AD___________________

Award Number: W81XWH-04-1-0559

TITLE: Concurrent MR-NIR Imaging for Breast Cancer Diagnosis

PRINCIPAL INVESTIGATOR: Birsen Yazici, Ph.D.

CONTRACTING ORGANIZATION: Rensselaer Polytechnic Institute
Troy, NY 12180-3590

REPORT DATE: June 2007

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
01-06-2007

2. REPORT TYPE
Annual

3. DATES COVERED *(From - To)*
1 Jun 2006 – 31 May 2007

4. TITLE AND SUBTITLE

Concurrent MR-NIR Imaging for Breast Cancer Diagnosis

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-04-1-0559

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Birsen Yazici, Ph.D.

E-Mail: yazici@ecse.rpi.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Rensselaer Polytechnic Institute
Troy, NY 12180-3590

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES – Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

The primary objective of this research program is to investigate concurrent near infrared (NIR) optical and magnetic resonance (MR) imaging for breast cancer diagnosis. The NIR diffuse optical imaging offers novel criteria for cancer differentiation with the ability to measure (in vivo) oxygenation and vascularization state, the uptake and release of contrast agents and chromophore concentrations with high sensitivity. However, NIR diffuse optical tomography is inherently a low spatial resolution imaging modality due to diffuse nature of light photons. Alternatively, MRI provides high spatial resolution with excellent tissue discrimination, but has limited ability to monitor hemoglobin dynamics and other contrast mechanisms that optical imaging provides. Therefore, concurrent MRI-NIR optical imaging brings together the most advantageous aspects of the two imaging modalities for breast cancer diagnosis.

15. SUBJECT TERMS
No subject terms provided.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	124	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Introduction……………………………………………………………..………...4-6

Body……………………………………………………………………….…….…6-30

Key Research Accomplishments……………………………………….………31

Reportable Outcomes…………………………………………………………...31-32

Conclusions……………………………………………………………….……...32-33

References…………………………………………………………………….…34-35

Appendices……………………………………………………………….………36-114

I. INTRODUCTION

Near infrared (NIR) diffuse optical imaging provides quantitative functional information from breast tissue that can not be obtained by conventional radiological methods. NIR techniques can provide *in vivo* measurements of oxygenation and vascularization state, the uptake and release of molecular contrast agents and chromophore concentrations with high sensitivity. There is considerable evidence that tumor growth is dependent on angiogenesis [1]- [3], and that tumor aggressiveness can be assessed from its increased number of new vessels and reduced oxygenation state relative to normal breast tissue and benign breast lesions [4]- [6]. NIR diffuse optical tomographic (DOT) methods has the potential to characterize angiogenesis related vessel density as it measures the total hemoglobin concentration and provide the ability to differentiate between benign and malignant lesions based on oxygen saturation. Furthermore, NIR methods are non-ionizing, relatively inexpensive and can be made portable.

The diagnosis and management of cancer involves several stages where magnetic resonance (MR) plays a valuable and growing role. MRI of the breast is now a routine part of the clinical care in many centers [9]- [11]. Magnetic Resonance imaging (MRI) is indicated in patients with inconclusive clinical and/or mammographic examinations. Patients that may benefit include women with radiographically dense breasts, and high risk potential population [12]- [13]. MRI possesses less than 10% contrast for soft tissue pathology [14]. Gadolinium (Gd) enhanced MRI offers much better contrast and is specific for tumor vessel imaging. However, the signal in the Gd-MRI arises from the larger vessels as the contrast agent is flushed out of the vascular bed of the tumor [15]. In comparison, NIR measurements of absorption have extremely high contrast. It was reported that 5% change in vascular density as measured histologically in ductal carcinomas leads to approximately 300% contrast in NIR absorption coefficients [7]. Furthermore, there are studies suggesting that the kinetics of contrast enhanced optical spectroscopy provides information about the cellular spaces [8]. On the other hand, NIR DOT suffers from poor spatial resolution and as such, it is unlikely that NIR imaging will be a stand-alone screening method in the general population. Therefore, we believe that the concurrent MR and NIR imaging brings together the most advantageous aspects of the two imaging modalities (structural and functional). In the future, we envision that this multimodality imaging approach will lead to high resolution hemoglobin tomography and comprehensive quantitative functional tissue characterization to differentiate malignant and benign tumors.

In this project, the clinical studies are performed using the novel MR-NIR hybrid time-resolved spectroscopy (TRS) imager and fast Indocynine Green (ICG) enhanced spectroscopic imager developed by Dr. Chance, a Co-PI of this proposal, at the University of Pennsylvania (UPenn), Biophysics Department, Diffuse Optical Imaging and Spectroscopy Laboratory.

The central hypothesis of this project is that the concurrent MR-NIR diffuse optical tomographic methods coupled with fast contrast enhanced NIR spectroscopic methods provide fundamentally new quantitative functional and structural information for breast cancer tumor characterization and detection. This new information can be obtained by novel modeling, analysis and data fusion methods from the tomographic, temporal and cellular-based contrast measurements, which exploit fast imaging techniques together with TRS tomographic methods. In this project, we investigate new methods for multi-modality high spatial resolution hemoglobin tomography, pharmacokinetic modeling of molecular contrast agents based on fast NIR spectroscopy and analysis of structural and functional information provided by MR and NIR imaging methods for breast cancer detection based on receiver operating characteristics methodology. Specific aims of the project are as follows:

- Aim 1: Utilize a priori anatomical information provided by MRI, to reconstruct 3D high resolution hemoglobin, water and lipid concentration, and oxygen saturation images directly from 6 wavelength time resolved optical measurements. Evaluate improvements in image reconstruction between that of stand-alone NIR and concurrent MR-NIR measurements using water and lipid images obtained from MRI.
- Aim 2: Develop a compartmentalized pharmacokinetic modeling of ICG, optical contrast agent, and extract quantitative parameters that can characterize tumor metabolism and angiogenesis. Compare ICG kinetics with the Gadolinium, MR contrast agent, kinetics and biopsy findings.
- Aim 3: Evaluate accuracy of breast cancer diagnosis based on the quantitative functional information extracted from stand-alone NIR system. This information includes hemoglobin, water and lipid concentration, optical scatter power and oxygen saturation images, and ICG pharmacokinetic parameters. Evaluate the added value

of ICG kinetic parameters in breast cancer diagnosis.

- Aim 4: Combine NIR based breast cancer diagnosis features with the systematic MR breast architecture and kinetics interpretation model developed by Dr. Nunes, M.D, Co-PI of this proposal, to evaluate the sensitivity and specificity of concurrent MR-NIR imaging method. Compare results with that of stand-alone MR and NIR results.

In the following sections, we will provide detailed description of our current research in line with the statement of work (SOW) and the aims outlined above. For the period of June 1st, 2006 to May 31st 2007, SOW includes only the first two aims of the project.

II. BODY

In the year 3 of the project, SOW includes tasks related to Aim 2 and 3 of the project. Below, we describe the works that has been performed towards the Aim 2, Task 3 and Aim 3 Tasks 3 and 4. We also improved upon the Task 1 of Aim 1.

A. AIM 2 - Tasks 3 / AIM 3- Task 1

1) Reconstruction of Bulk ICG Concentration Images: In our data collection process, a sequence of boundary measurements are collected over a period of time. Each set of measurements are used to form a frame of the ICG concentration images. The resulting sequence of ICG concentration images are then used to form pharmacokinetic-rate images. To reconstruct each frame of the ICG concentration images, we follow a static reconstruction approach and use differential diffuse optical tomography (DDOT) technique [18], [23].

In DDOT, two sets of excitation measurements are collected corresponding to before and after the ICG injection, and the ICG concentration is determined by the perturbation method [18], [23]. The photon propagation before and after the injection is modeled by the following diffusion equation:

$$\nabla \cdot [D_x \nabla \Phi_x^{\pm}(r,\omega) - [\mu_{ax}^{\pm}(r) + j\omega/c]\Phi_x^{\pm}(r,\omega) = 0, \qquad r \in \Omega \subseteq \mathbb{R}^3 \tag{1}$$

with Robin-type boundary conditions:

$$2D_x \frac{\partial \Phi_x^{\pm}}{\partial \nu} + \rho \Phi_x^{\pm} = -S(r,\omega), \qquad r \in \partial\Omega. \tag{2}$$

where x stands for the excitation, c is the speed of light inside the medium Ω; ω denotes the modulation frequency of the source, $\mu_{ax}^{-}(r)$ and $\mu_{ax}^{+}(r)$ are the absorption coefficients before and after the ICG injection, D_x is the diffusion coefficient which is assumed to be constant, $\Phi_x^{\pm}(r,\omega)$ optical field at location r before and after the ICG injection. Here, ν denotes the outward normal to the boundary $\partial\Omega$ of Ω, ρ is a constant representing the refractive index mismatch between the two regions separated by $\partial\Omega$, and $S(r,\omega)$ is the excitation source on the boundary.

The absorption coefficient after the injection μ_{ax}^{+} are modeled as a sum of the absorption coefficient of the medium before the ICG injection μ_{ax}^{-} and the perturbation caused by the ICG $\Delta\mu_{ax}(r)$:

$$\Delta\mu_{ax}(r) = \mu_{ax}^{+}(r) - \mu_{ax}^{-}(r), \qquad r \ \in \ \Omega \ \subseteq \ \mathbb{R}^3. \tag{3}$$

In the forward model, the analytical solutions of the heterogonous diffusion equation given in (1) is derived using first order Rytov approximation [18]. The sample volume is divided into a set of voxels and the measurements are related to the relative absorption coefficients of each voxel by a system of linear equations. The shape of the breast was approximated as a cylinder and the Kirchhoff approximation [24], [25] for diffuse waves was used to model the interaction of light with boundaries. In order to account for the biological noise, the forward model was implemented with coupling coefficient technique [26].

Here, the Rytov-type measurements, which are defined by the natural logarithm of the ratio of the post-ICG measurements to the pre-ICG measurements were used [23]. Let $\Psi_x(r_d,\omega;r_s)$ denote the Rytov-type measurements at location r_d due to source at r_s. The linearized relationship between the differential absorption coefficient and measurements is given by [21],

$$\Psi_x(r_d,\omega;r_s) = -\frac{1}{\Phi_x^{-}(r_d,\omega;r_s)}\int_{\Omega} G_x^{-}(r-r_d,\omega;r_s) I_x(r)\Phi_x^{-}(r,\omega;r_s)d^3r \tag{4}$$

where $\Phi_x^{-}(r,\omega;r_s)$ is the photon density obtained at the excitation wavelength before ICG injection, $I_x(r) = c\Delta\mu_{ax}(r)/D_x$, and $G_x^{-}(r-r_d,\omega;r_s)$ is the Green's function of (1) for a source at r_s before the injection describing the propagation of light from the heterogeneity r to the detector at r_d.

We address the inverse problem of recovering $\Delta\mu_{ax}$ from Rytov-measurements Ψ_x based on the forward model (4) using the singular value decomposition of the Moore-Penrose generalized system. We use a zeroth-order Tikhonov regularization to stabilize the inversion procedure. We apply the L-curve method to an experimental model reconstruction and derive the best regularization parameter using a curvature function as described in [27]. A detailed discussion of the forward and inverse models used for the reconstruction of differential absorption coefficients ($\Delta\mu_{ax}$) can be found in [18].

To construct a set of ICG concentration images, we use the linear relationship between the differential absorption coefficients and ICG concentrations [20]:

$$\Delta\mu_a(r) = \ln 10\ \epsilon_\lambda\ m(r)\ = 2.303\ \epsilon_\lambda\ m(r) \tag{5}$$

where ϵ_λ is the extinction coefficient of ICG at the wavelength 805nm, $m(r)$ is the bulk ICG concentration in the tissue, and $\Delta\mu_a(r)$ is as defined in (3).

2) Modeling and Estimation of ICG Pharmacokinetics: Two-compartment Model of ICG Pharmacokinetics: Using the method outlined in Section 2, we reconstruct a sequence of ICG concentration images. As an example, Figure 3 shows a set of images reconstructed from *in vivo* breast data.

Our objective is to model the pharmacokinetics of ICG at each voxel of ICG concentration images using compartmental modeling. To do so, we first extracted the time varying ICG concentration curves for each voxel from the sequence of ICG concentration images. An example of such a curve is shown in Figure 6. We next fit a two-compartment model to each ICG concentration curve [19], [22]. Figure 23 shows the two-compartment model for ICG kinetics. Here, plasma and extracellular-extravascular space (EES) constitute the two compartments. C_p and C_e represent the ICG concentrations in plasma and the EES, respectively. The rates k_{in} and k_{out} govern the leakage into and the drainage out of the EES. The parameter k_{elm} describes the ICG elimination from the body through kidneys and livers.

Using the two-compartment model introduced in [19], ICG transition between plasma and the EES can be modeled as follows:

$$\begin{bmatrix} \dot{C}_e(t) \\ \dot{C}_p(t) \end{bmatrix} = \begin{bmatrix} -k_{out} & k_{in} \\ k_{out} & -(k_{in}+k_{elm}) \end{bmatrix} \begin{bmatrix} C_e(t) \\ C_p(t) \end{bmatrix} + \boldsymbol{\omega}(t), \tag{6}$$

where $C_p(t)$ and $C_e(t)$ represent the ICG concentrations in plasma and the EES at time $t \in [T_0, T_1]$, respectively. The rates k_{in}, k_{out}, and k_{elm} have a unit of sec^{-1}. They are defined as the permeability surface area products given by $PS\rho$, where P is the capillary permeability constant, S is the capillary surface area, and ρ is the tissue density. Here, $\boldsymbol{\omega}(t)$ is uncorrelated zero-mean Gaussian process with covariance matrix **Q** representing the model mismatch.

The actual total ICG concentration in the tissue is a linear combination of plasma and the EES ICG concentrations modeled as:

$$m(t) = \begin{bmatrix} v_e & v_p \end{bmatrix} \begin{bmatrix} C_e(t) \\ C_p(t) \end{bmatrix} + \eta(t) \tag{7}$$

where $m(t)$, $C_e(t)$, and $C_p(t)$ are defined in (5) and (6); v_p and v_e are plasma and the EES volume fractions, respectively; and $\eta(t)$ is uncorrelated zero-mean Gaussian process with covariance matrix $\mathbf{R}$, representing the measurement noise.

Estimation of ICG Pharmacokinetics using Extended Kalman Filtering: In matrix-vector notation, (6) and (7) can be expressed as:

$$\dot{\mathbf{C}}(t) = \mathbf{K}(\boldsymbol{\alpha})\mathbf{C}(t) + \boldsymbol{\omega}(t), \tag{8}$$

$$m(t) = \mathbf{V}(\boldsymbol{\alpha})\mathbf{C}(t) + \eta(t),$$

where $\mathbf{C}(t)$ denotes the concentration vector with elements $C_e(t)$, and $C_p(t)$; $\mathbf{K}(\boldsymbol{\alpha})$ is the system matrix, $\mathbf{V}(\boldsymbol{\alpha})$ is the measurement matrix as defined in equation (7), and $\boldsymbol{\alpha}$ is the parameter vector given by

$$\boldsymbol{\alpha} = [k_{out} \;\; k_{in} \;\; k_{elm} \;\; v_e \;\; v_p]^T. \tag{9}$$

The ICG measurements in (8) are collected at discrete time instances, $t = kT$, $k = 0, 1, ...$, where T is the sampling period. Therefore, the continuous model described in (8) is discretized. We can express the discrete compartmental model as follows:

$$\mathbf{C}_d(k+1) = \mathbf{K}_d(\boldsymbol{\theta})\mathbf{C}_d(k) + \boldsymbol{\omega}_d(k), \tag{10}$$

$$\mathbf{m}(k) = \mathbf{V}_d(\boldsymbol{\theta})\mathbf{C}_d(k) + \boldsymbol{\eta}_d(k),$$

where $\mathbf{K}_d(\boldsymbol{\theta}) = e^{\mathbf{K}(\boldsymbol{\alpha})}$ is the discrete time system matrix; $\mathbf{V}_d(\theta) = \mathbf{V}(\boldsymbol{\alpha})$ is the discrete measurement matrix; $\boldsymbol{\omega}_d(k)$ and $\boldsymbol{\eta}_d(k)$ are zero-mean Gaussian white noise processes with covariances matrix $\mathbf{Q}_d$ and variance $\mathbf{R}_d$, respectively. The vector $\boldsymbol{\theta}$ is composed of parameters τ_{ij} which are functions the pharmacokinetic-rates and volume fractions:

$$\boldsymbol{\theta} = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{21} & \tau_{22} & v_e & v_p \end{bmatrix}^T. \tag{11}$$

We first estimate τ_{ij}'s, $i, j = 1, 2$ and then compute the pharmacokinetic-rates k_{in}, k_{out} and k_{elm} [19], [28]. The explicit form of the discrete state-space model is given as follows:

$$\begin{bmatrix} C_e(k+1) \\ C_p(k+1) \end{bmatrix} = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{bmatrix} \begin{bmatrix} C_e(k) \\ C_p(k) \end{bmatrix} + \boldsymbol{\omega}_d(k) \tag{12}$$

$$m(k) = \begin{bmatrix} v_e & v_p \end{bmatrix} \begin{bmatrix} C_e(k) \\ C_p(k) \end{bmatrix} + \eta_d(k).$$

We estimate the parameter vector $\boldsymbol{\theta}$ and concentration vector $\mathbf{C}_d$ by using the EKF framework. The EKF is a recursive modeling and estimation method with numerous advantages in ICG pharmacokinetic modeling [19].

These include effective modeling of multiple compartments, and multiple measurement systems in the presence of measurement noise and uncertainties in the compartmental model dynamics, simultaneous estimation of model parameters and ICG concentrations in each compartment, statistical validation of estimated concentrations and error bounds on the model parameter estimates, and incorporation of available a priori information about the initial conditions of the permeability rates into the estimation procedure.

When both states (ICG concentrations) and model parameters (pharmacokinetic-rates and volume fractions) are unknown, a linear state-space model can be regarded as a non-linear model; the linear system parameters and states combine to form the new states of the non-linear model. This system is then linearized and the new unknown states are found using the EKF estimator [19], [29]–[31]. In EKF framework, $\boldsymbol{\theta}$ can be treated as a random process with the following model:

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \varsigma_d(k), \tag{13}$$

where $\varsigma_d(k)$ is a zero-mean Gaussian process with covariance matrix $\mathbf{S}_d$.

Table 1 summarizes the joint estimation of pharmacokinetic-rates and ICG concentration in different compartments. In Table 1, $\hat{\mathbf{C}}_d(k|k-1)$ is the state estimate propagation at step k given all the measurements up to step $k-1$; $\hat{\mathbf{C}}_d(k)$ is the state estimate update at step k; $\mathbf{P}_{k,k-1}$ denotes the error covariance propagation at step k given all the measurements up to step $k-1$; $\mathbf{P}_{k,k}$ is the error covariance update at step k; $\mathbf{S}_d$ is the preassigned covariance matrix of $\varsigma_d(k)$; $\mathbf{J}_k$ is the Jacobian matrix due to iterative linearization of the state equation at step k; $\mathbf{G}_k$ is the recursive Kalman gain at step k; $\mathbf{R}_d$ is the covariance matrix of the measurements; $\mathbf{Q}_d$ is the covariance matrix of the concentration vector; and $\mathbf{I}$ is the identity matrix. A detailed discussion of the Kalman filtering algorithm for the joint estimation of the model parameters and ICG concentrations in different compartments can be found in [19].

3) Clinical Results - Apparatus: : In this work, we use the data collected with a continuous wave (CW) NIR imaging apparatus. The apparatus has 16 light sources, which are tungsten bulbs with less than 1 watt of output power. They are located on a circular holder at an equal distance from each other with 22.5 degrees apart. Sixteen detectors, namely, silicon photodiodes, are situated in the same plane. The breast is arranged in a pendular geometry with the source-detector probes gently touching its surface. Figure 2 illustrates the configuration of the apparatus and the configuration of the detectors and the sources in a circular plane. Note that sources and detectors are co-located. The detectors use the same positions as the sources to collect the light originating from one source at a time. Only the signals from the farthest 11 detectors are used in the analysis. For example, when Source 1 is on, the data is collected using Detectors 4 to 14. A band pass filter at 805 nm, the absorption peak of ICG, is placed in front of the sources to select the desired wavelength. A set of data for one source is collected every $\sim$ 500 ms. The total time for a whole scan of the breast including 16 sources and 16 detectors is $\sim$ 8.8 seconds. A more detailed explanation of the apparatus and the data collection procedure can be found in [32].

Protocol: Patients with suspicious breast tumors are enrolled for this study. ICG is injected intravenously by bolus with a concentration of 0.25 mg per kg of body weight. Diagnostic information is obtained using biopsy results. Since biopsy modifies the blood volume and blood flow around the tumor region, measurements are made prior to biopsy. Data acquisition started before the injection of ICG and continued for 10 minutes.

Tumor Information: Three different patients with different tumor types are included in this study. First case, (Case 1), is fibroadenoma, which corresponds to a mass estimated to be $1-2$ cm in diameter within a breast of 9 cm diameter located at 6-7 o'clock. Second case, (Case 2), is adenocarcinoma corresponding to a tumor estimated to be $2-3$ cm in diameter within a breast of 7.7 cm diameter located at 4-5 o'clock. The third case, (Case 3), is invasive ductal carcinoma, which corresponds to a mass estimated to be 4 by 3 cm located at 6 o'clock. Table 2 describes the tumor information for each patient.

4) ICG Concentration Measurements for Pharmacokinetic Parameter Estimations: Using the CW imager described above, sufficient number of source-detector readings are collected from different angles for each patient. Differential absorption images are reconstructed based on DDOT forward model given in (4). Using the linear relationship (5) between ICG concentration, and absorption coefficient, ICG concentration images are obtained for each case. A sample set of ICG concentration images for the selected time instants are shown in Figure 3, 4, and 5 for Case 1, 2, and 3, respectively. Although only 9 images are displayed, there are around 50 images for each case, each corresponding to a different time instant. Each image is composed of 649 voxels. Note that the ICG concentration images in Figures 3, 4, and 5 represent the bulk ICG concentrations in the tissue, not the ICG concentrations in plasma or in the EES compartments.

We next extract the time course of ICG concentration for each voxel. As an example, Figure 6 shows the time course of ICG concentrations for all three cases for a specific voxel around the tumor region (65th, 276th, 188th voxel for Case 1, Case 2, and Case 3, respectively). We then fit the two-compartment model to each time course data using the EKF framework; and estimate k_{in}, k_{out}, k_{elm}, and the ICG concentrations in plasma and EES. Spatially resolved images of k_{in}, and k_{out} for each case are shown in Figures 8(a)-(b), and 9(a)-(b), 10(a)-(b), respectively. Additionally, we construct ICG concentration images for plasma and the EES compartments. Figures

11-16 show the ICG concentration in plasma and the EES for 3 different time instants for Case 1, 2, and 3, respectively. Our results show that the pharmacokinetic-rates are higher around the tumor region agreeing with the fact that permeability increases around the tumor region due to compromised capillaries of cancer vessels. We also observed that ICG concentrations in plasma and the EES compartments are higher around the tumors agreeing with the hypothesis that around the tumor region ICG may act as a diffusible extravascular flow in leaky capillary of cancer vessels.

Using the brightest spot as a reference point, the *a priori* information on the location, and the size of the tumors, we plot an ellipse (or a circle) to identify the approximate location and size of the tumor in the pharmacokinetic-rate images. Figures 17(a), (b), and (c) present the k_{in} images with approximate tumor location and size for Case 1, 2, and 3, respectively. The consistency of the bright regions in the k_{in} images, and circular/elliptical regions drawn based on the *a priori* information shows that the pharmacokinetic-rate images provide good localization of tumors and size information.

The histograms of k_{in} and k_{out} rates from inside and outside the tumor region (as indicated by circular/elliptical regions) are shown in Figures 18(a), (b), and (c), and Figures 19(a), (b), and (c), respectively. The solid curves in Figures 18 and 19 show the Gaussian fit. Clearly, pharmacokinetic-rate images facilitates analysis of the spatial variation of the pharmacokinetic-rates for breast cancer screening, diagnosis, and staging. The histograms and their Gaussian probability density function fits in Figures 18 and 19 show that the mean and the standard deviation of k_{in} and k_{out} values are different inside and outside the tumor region. Table 3 tabulates the mean values (± spatial standard deviation) of the pharmacokinetic-rates from inside and outside the tumor regions for all three cases. The total error (probability of false positive and false negative) using the Bayesian minimum error classifier is given by the overlapping area between the two Gaussian fits [33]. Table 4 tabulates the total Bayesian error. These small error values show that the spatially resolved pharmacokinetics-rate analysis provide good sensitivity and specificity for diagnosis. One other advantage of the pharmacokinetic-rate images is that they allow comparison of rates obtained from healthy and tumor tissue for the patient itself since absolute (mean) values may be quite different from patient to patient. For example, the k_{in} rate for Case 3 obtained outside the tumor region is higher than the k_{in} rate for Case 1 obtained inside the tumor. Thus, using the pharmacokinetic-rates obtained from healthy tissue as a reference to the tumor tissue for the same patient can improve diagnostic accuracy. However, very high mean values of k_{in} and k_{out} may be indicative of the severity of malignancy. The pharmacokinetic rates are higher for Case 3 (invasive ductal carcinoma), for both inside and outside the tumor region as compared to Case 2 (adenocarcinoma). Similarly, the kinetic rates are higher for Case 2 (adenocarcinoma), as compared to Case 1 (fibroadenoma) for both inside and outside the tumor region.

To further understand the value of spatially resolved pharmacokinetic-rate analysis as compared to the bulk pharmacokinetic-rates, we averaged the concentration images spatially, and obtained a bulk concentration value for each time instant. We then extracted a time curve for the bulk ICG concentrations. Figure 7 shows the resulting bulk ICG concentration time curves for each case. Next, we fitted the two-compartment model to the resulting time curves and estimated the bulk pharmacokinetic-rates. Table 5 tabulates the bulk pharmacokinetic-rates for each patient. In Figure 20 and 21, the bulk pharmacokinetic-rates are overlaid on the histograms of the pharmacokinetic-rate images obtained from inside and outside the tumor region. Figures 20 and 21 show the Bayesian minimum error classifier threshold (the value corresponding to the intersection of the histograms) for each case. We see that for Case 1, the bulk rates of k_{in}, and k_{out} are both classified as healthy tissue (outside the tumor region). For Case 2, k_{in} is classified as cancerous tissue (inside the tumor region) and k_{out} is classified as healthy tissue. Similarly for Case 3, k_{in} is classified as healthy and k_{out} is classified as cancerous tissue. Clearly, the bulk rates do not provide consistent diagnosis. On the other hand, spatially resolved pharmacokinetic-rates allow us to compare values from the cancerous region to the healthy background, and provide higher sensitivity, and specificity as shown by the Bayesian total error values in Table IV.

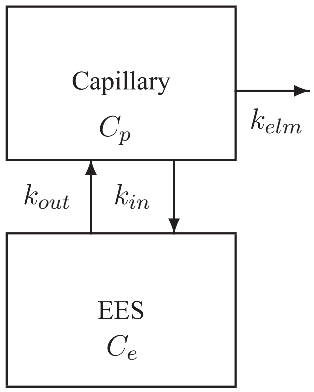


Fig. 1. Block diagram of the two-compartment model for ICG pharmacokinetics.

TABLE I

EKF algorithm for simultaneous estimation of states and parameters.

Initial Conditions	$\begin{bmatrix} \hat{\mathbf{C}}_d(0) \\ \hat{\boldsymbol{\theta}}(0) \end{bmatrix} = \begin{bmatrix} E(\mathbf{C}_d(0)) \\ \hat{\boldsymbol{\theta}}(0) \end{bmatrix}, \quad \mathbf{P}_{0,0} = \begin{bmatrix} Var(\mathbf{C}_d(0)) & 0 \\ 0 & \mathbf{S}_d \end{bmatrix}$
State Estimate Propagation	$\begin{bmatrix} \hat{\mathbf{C}}_d(k\|k-1) \\ \hat{\boldsymbol{\theta}}(k\|k-1) \end{bmatrix} = \begin{bmatrix} \mathbf{K}_d(\hat{\boldsymbol{\theta}}(k-1))\hat{\mathbf{C}}_d(k-1) \\ \hat{\boldsymbol{\theta}}(k-1) \end{bmatrix}$
Error Covariance Propagation	$\mathbf{P}_{k,k-1} = \mathbf{J}_{k-1}\mathbf{P}_{k-1,k-1}\mathbf{J}_{k-1}^T + \begin{bmatrix} \mathbf{Q}_d & 0 \\ 0 & \mathbf{S}_d \end{bmatrix}$
State Estimate Update	$\begin{bmatrix} \hat{\mathbf{C}}_d(k) \\ \hat{\boldsymbol{\theta}}(k) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{C}}_d(k\|k-1) \\ \hat{\boldsymbol{\theta}}(k\|k-1) \end{bmatrix}$ $+\mathbf{G}_k(\mathbf{m}(k) - \mathbf{V}_d(\boldsymbol{\theta}(k\|k-1))\mathbf{C}_d(k\|k-1))$
Error Covariance Update	$\mathbf{P}_{k,k} = [\mathbf{I} - \mathbf{H}_k\boldsymbol{\Lambda}_{k\|k-1}]\mathbf{P}_{k,k-1}$
Kalman Gain	$\mathbf{H}_k = \mathbf{P}_{k,k-1}\boldsymbol{\Lambda}_{k\|k-1}^T[\boldsymbol{\Lambda}_{k\|k-1}\mathbf{P}_{k,k-1}\boldsymbol{\Lambda}_{k\|k-1}^T + \mathbf{R}_d]^{-1}$
Definitions	$\mathbf{J}_k = \begin{bmatrix} \mathbf{K}_d(\hat{\boldsymbol{\theta}}(k)) & \frac{\partial}{\partial\boldsymbol{\theta}}[\mathbf{K}_d(\hat{\boldsymbol{\theta}}(k))\hat{\mathbf{C}}_d(k)] \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \boldsymbol{\Lambda}_{k\|k-1} = \begin{bmatrix} \mathbf{V}_d(\boldsymbol{\theta}(k\|k-1)) \\ 0 \end{bmatrix}^T$

TABLE II

Tumor information for each patient

	Tumor Type	Tumor Size	Tumor Location
Case 1	Fibroadenoma	1-2 cm	6-7 o‘clock
Case 2	Adenocarcinoma	2-3 cm	4-5 o‘clock
Case 3	Invasive Ductal Carcinoma	4 by 3 cm	6 o‘clock

TABLE III

Mean and standard deviation of pharmacokinetic-rates for inside and outside the tumor region

	k_{in} $(sec^{-1}10^{-2})$		k_{out} $(sec^{-1}10^{-2})$		k_{elm} $(sec^{-1}10^{-3})$	
	Inside	Outside	Inside	Outside	Inside	Outside
Case 1	2.14±0.018	0.73±0.011	1.24±0.069	0.43±0.013	4.11±0.057	3.87±0.012
Case 2	2.92±0.076	1.14±0.052	1.58±0.051	0.65±0.036	3.94±0.081	4.12±0.047
Case 3	6.87±0.093	3.06±0.015	4.96±0.048	1.66±0.072	4.49±0.056	4.46±0.081

TABLE IV

TOTAL ERROR (PROBABILITY OF FALSE POSITIVE AND FALSE NEGATIVE) IN TUMOR DIAGNOSIS BASED ON BAYESIAN MINIMUM ERROR CLASSIFIER

	Total Error for k_{in}	Total Error for k_{out}
Case 1	6.67×10^{-2}	13.34×10^{-2}
Case 2	3.23×10^{-2}	13.51×10^{-2}
Case 3	9.09×10^{-2}	3.64×10^{-2}

TABLE V

BULK PHARMACOKINETIC-RATES EXTRACTED FROM THE ENTIRE BREAST TISSUE

	k_{in} $(sec^{-1}10^{-2})$	k_{out} $(sec^{-1}10^{-2})$	k_{elm} $(sec^{-1}10^{-1})$
Case 1	0.84±0.013	0.62±0.017	3.66±0.042
Case 2	2.01±0.022	0.83±0.012	4.01±0.054
Case 3	4.06±0.072	3.36±0.051	4.37±0.052

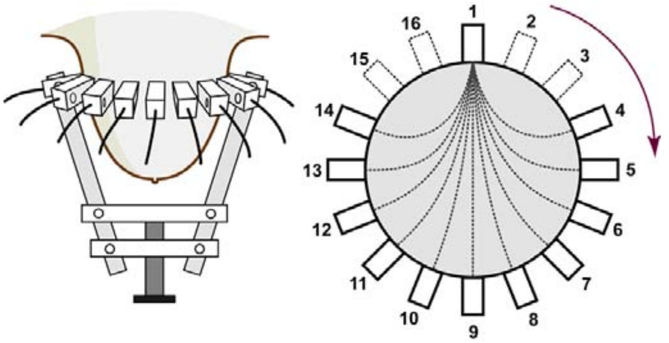


Fig. 2. The cut section of the CW NIR imaging apparatus with 16 sources and detectors.

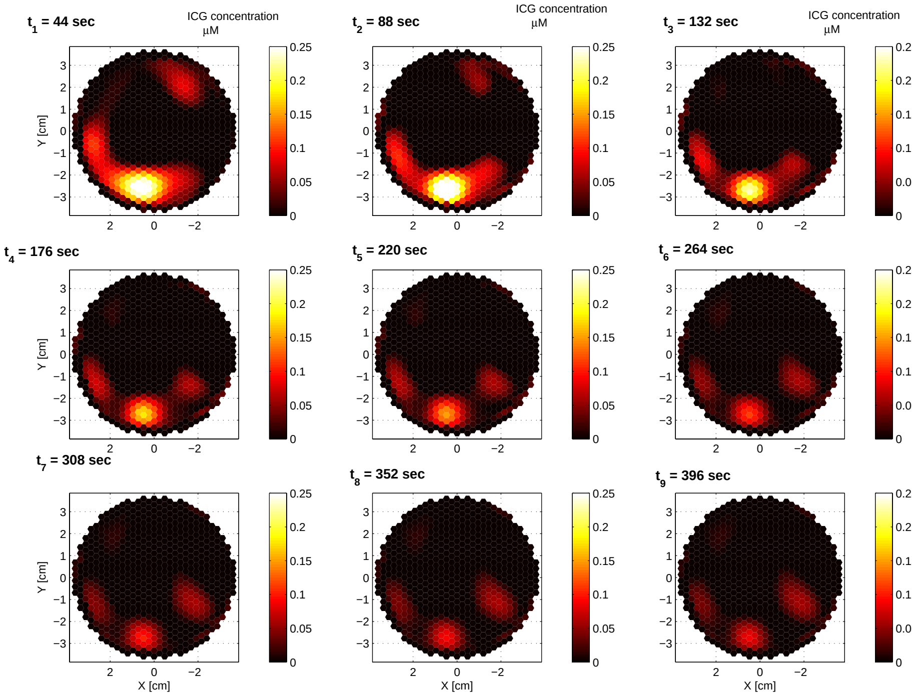


Fig. 3. Spatially resolved bulk ICG concentration images for a set of time instants for Case 1.

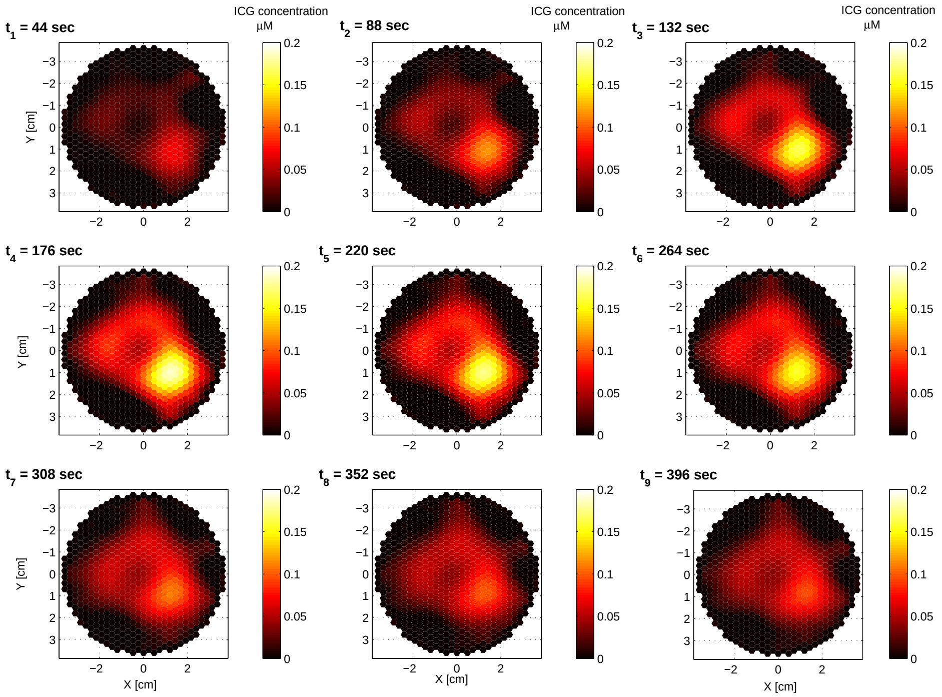


Fig. 4. Spatially resolved bulk ICG concentration images for a set of time instants for Case 2.

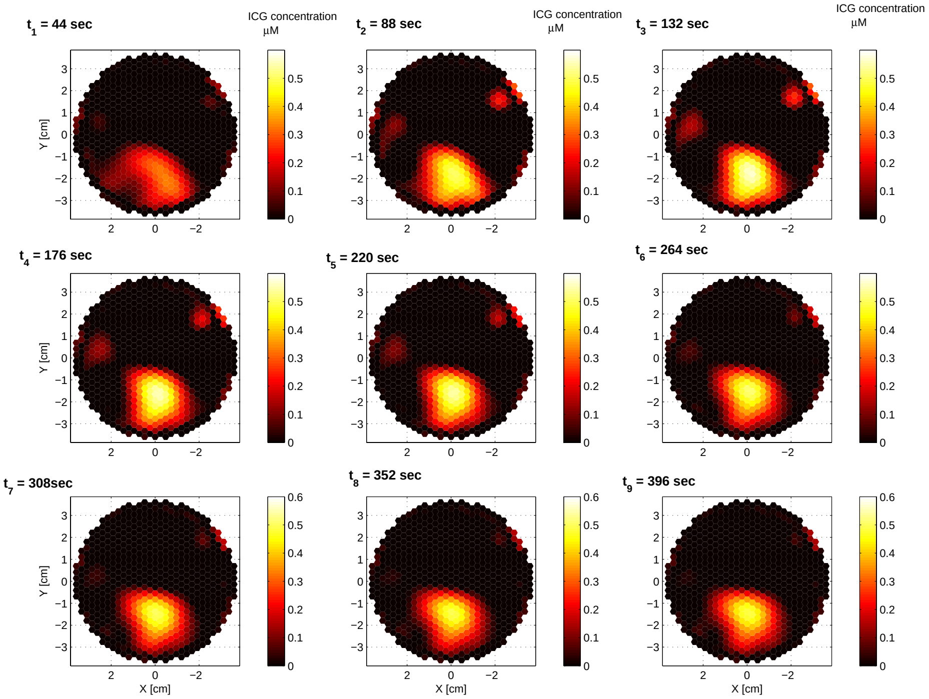


Fig. 5. Spatially resolved bulk ICG concentration images for a set of time instants for Case 3.

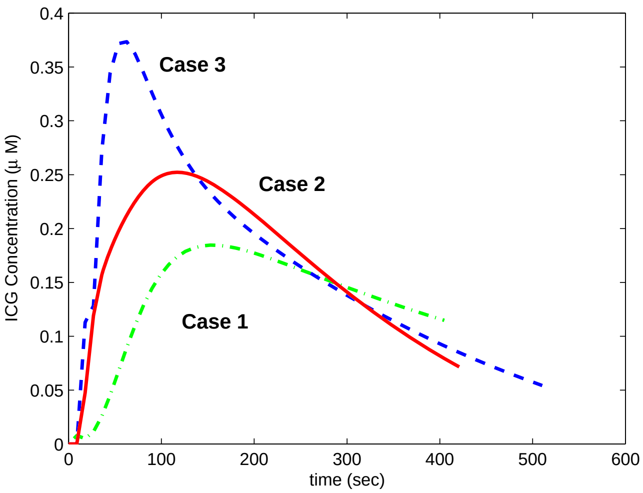


Fig. 6. Time course of ICG concentration curves for a specific voxel, 65th, 276th, 188th voxel for Case 1, Case 2, and Case 3, respectively.

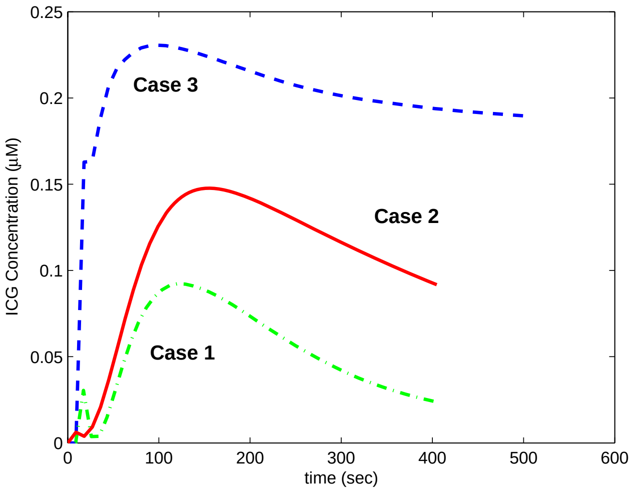
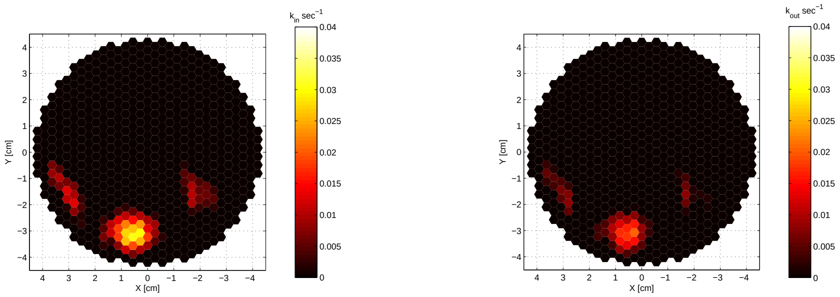


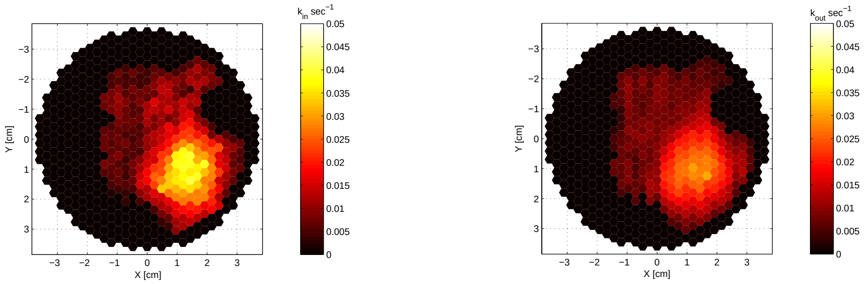
Fig. 7. Averaged time course of ICG concentration curves for Case 1, Case 2, and Case 3.


k_in sec^−1
k_out sec^−1
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
Y [cm]
X [cm]
4
3
2
1
0
−1
−2
−3
−4

(a)

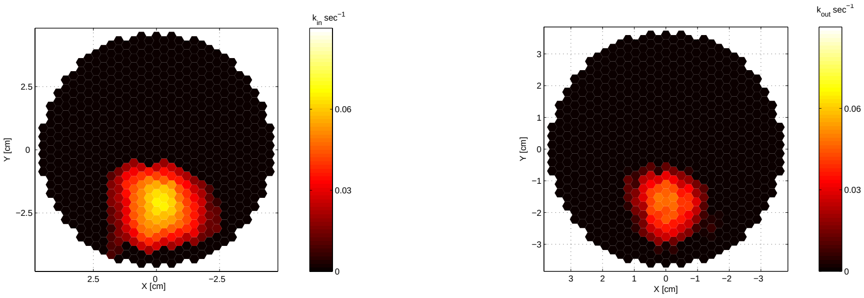
(b)

Fig. 8. Spatially resolved pharmacokinetic-rate images, (a) k_{in}, and (b) k_{out} for Case 1.


k_in sec^-1
k_out sec^-1
Y [cm]
X [cm]
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0

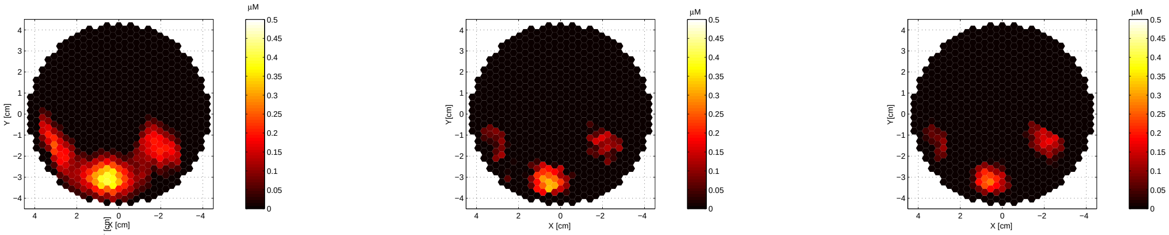
(a) (b)

Fig. 9. Spatially resolved pharmacokinetic-rate images, (a) k_{in}, and (b) k_{out} for Case 2.


k_in sec^-1
k_out sec^-1
Y [cm]
X [cm]
0.06
0.03
0

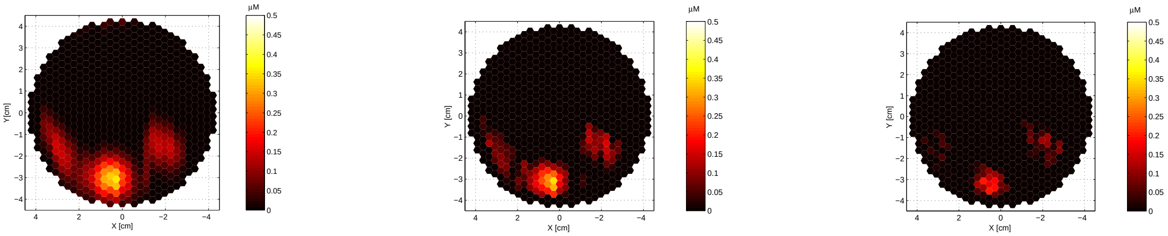
(a) (b)

Fig. 10. Spatially resolved pharmacokinetic-rate images, (a) k_{in}, and (b) k_{out} for Case 3.


µM
Y [cm]
X [cm]
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0

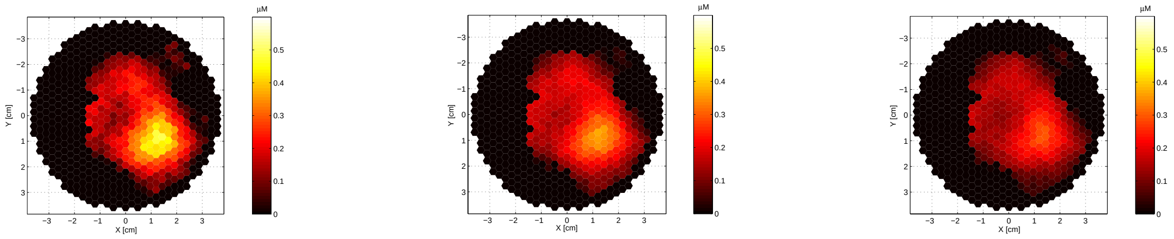
(a) (b) (c)

Fig. 11. Spatially resolved ICG concentration images in plasma for Case 1 for the 246.4^{th}, 334.4^{th}, and 422.4^{th} seconds.


µM
Y [cm]
X [cm]
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0

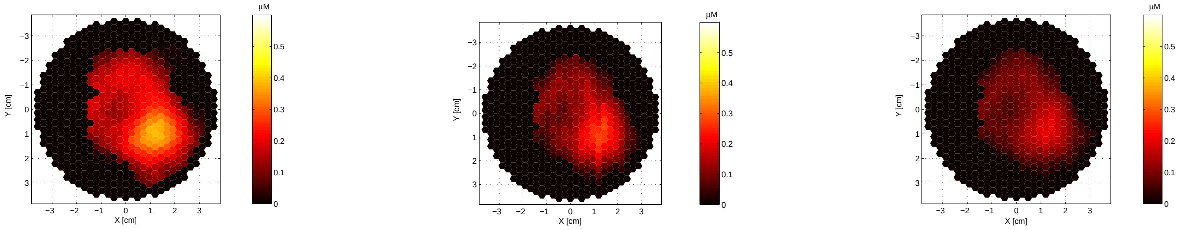
(a) (b) (c)

Fig. 12. Spatially resolved ICG concentration images in the EES for Case 1 for the 246.4^{th}, 334.4^{th}, and 422.4^{th} seconds.


Y [cm]
X [cm]
μM
−3
−2
−1
0
1
2
3
0
0.1
0.2
0.3
0.4
0.5

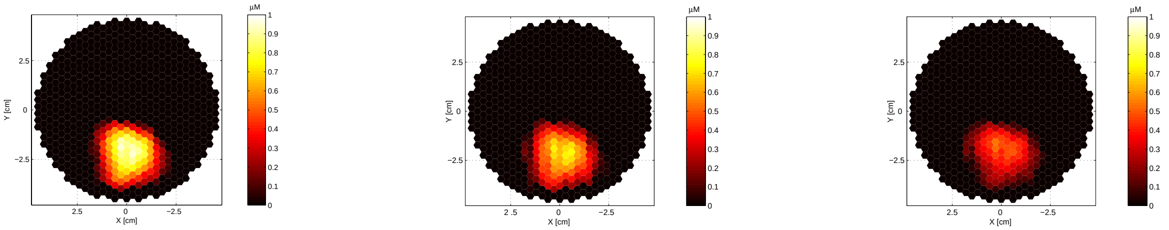
(a) (b) (c)

Fig. 13. Spatially resolved ICG concentration images in plasma for Case 2 for the 228.8^{th}, 316.8^{th}, and 404.8^{th} seconds.


Y [cm]
X [cm]
μM
−3
−2
−1
0
1
2
3
0
0.1
0.2
0.3
0.4
0.5

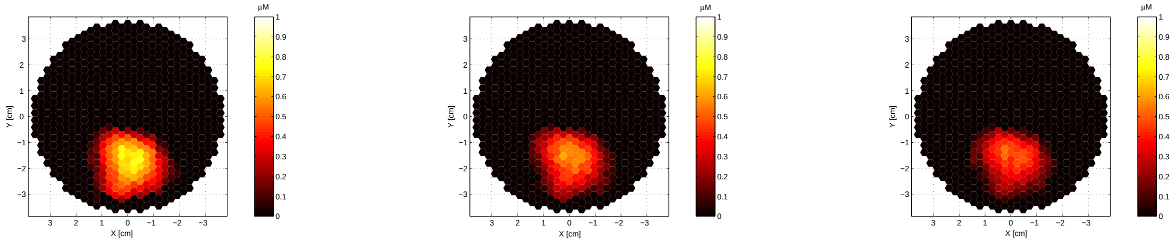
(a) (b) (c)

Fig. 14. Spatially resolved ICG concentration images in the EES for Case 2 for the 228.8^{th}, 316.8^{th}, and 404.8^{th} seconds.


Y [cm]
X [cm]
μM
2.5
0
−2.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

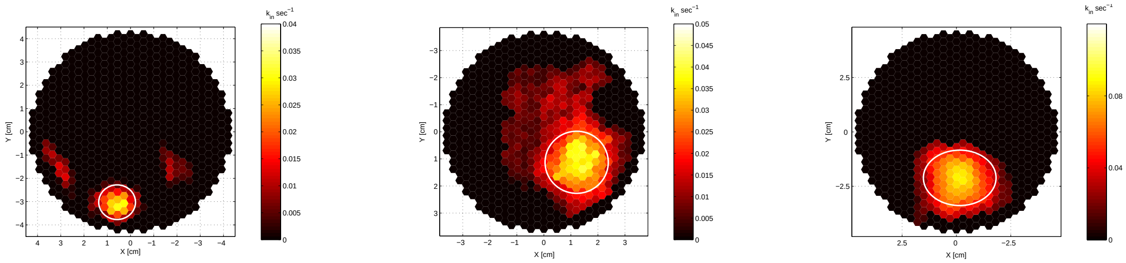
(a) (b) (c)

Fig. 15. Spatially resolved ICG concentration images in plasma for Case 3 for the 246.4^{th}, 378.4^{th}, and 510.4^{th} seconds.


Y [cm]
X [cm]
μM
3
2
1
0
−1
−2
−3
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

(a) (b) (c)

Fig. 16. Spatially resolved ICG concentration images in the EES for Case 3 for the 246.4^{th}, 378.4^{th}, and 510.4^{th} seconds.

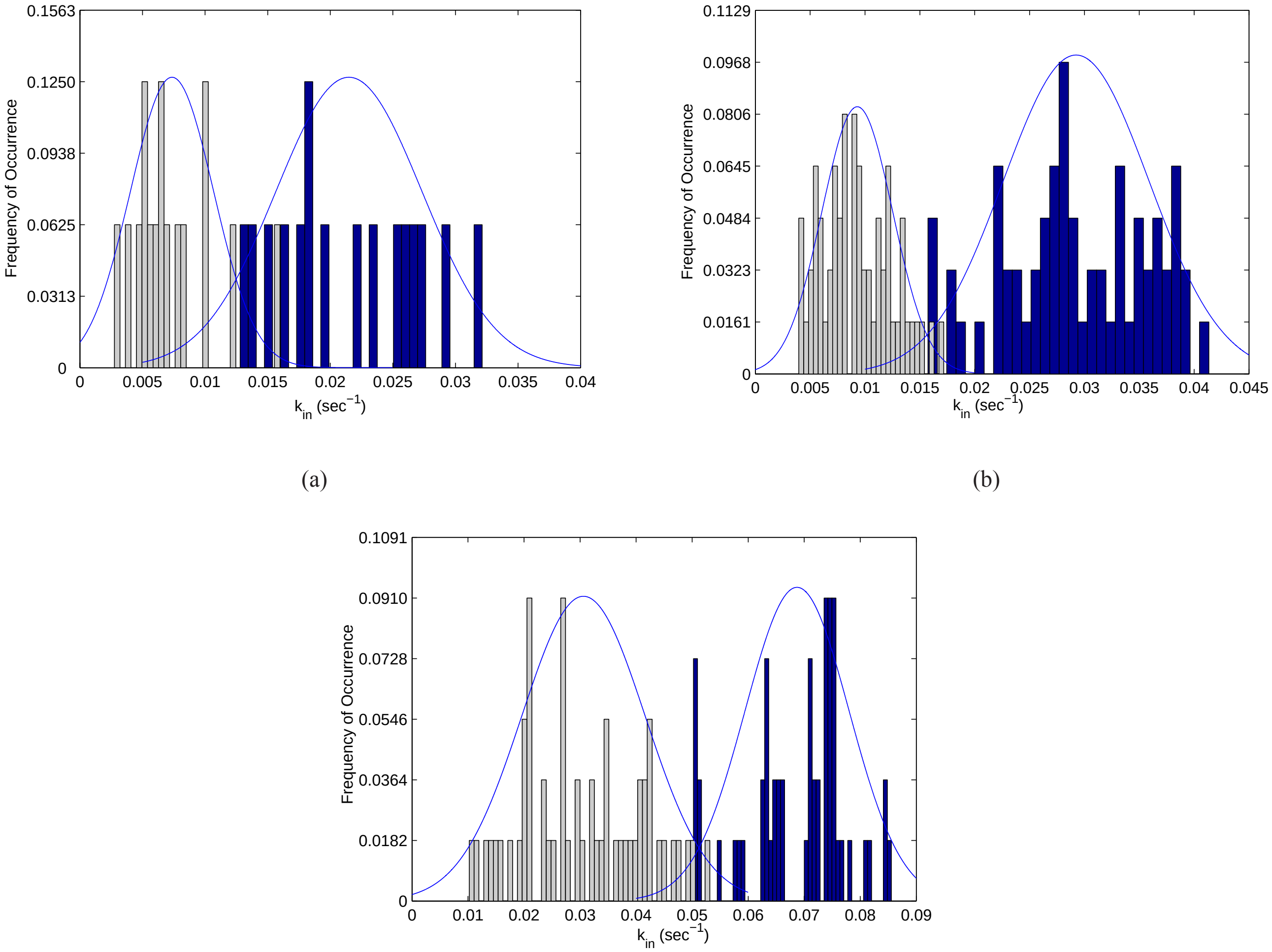


(a)

(b)

(c)

Fig. 17. The k_{in} images with approximate tumor location and size for Case 1, Case 2, and Case 3, respectively. An ellipse or a circle identifies the approximate location, shape, and size of the tumor.

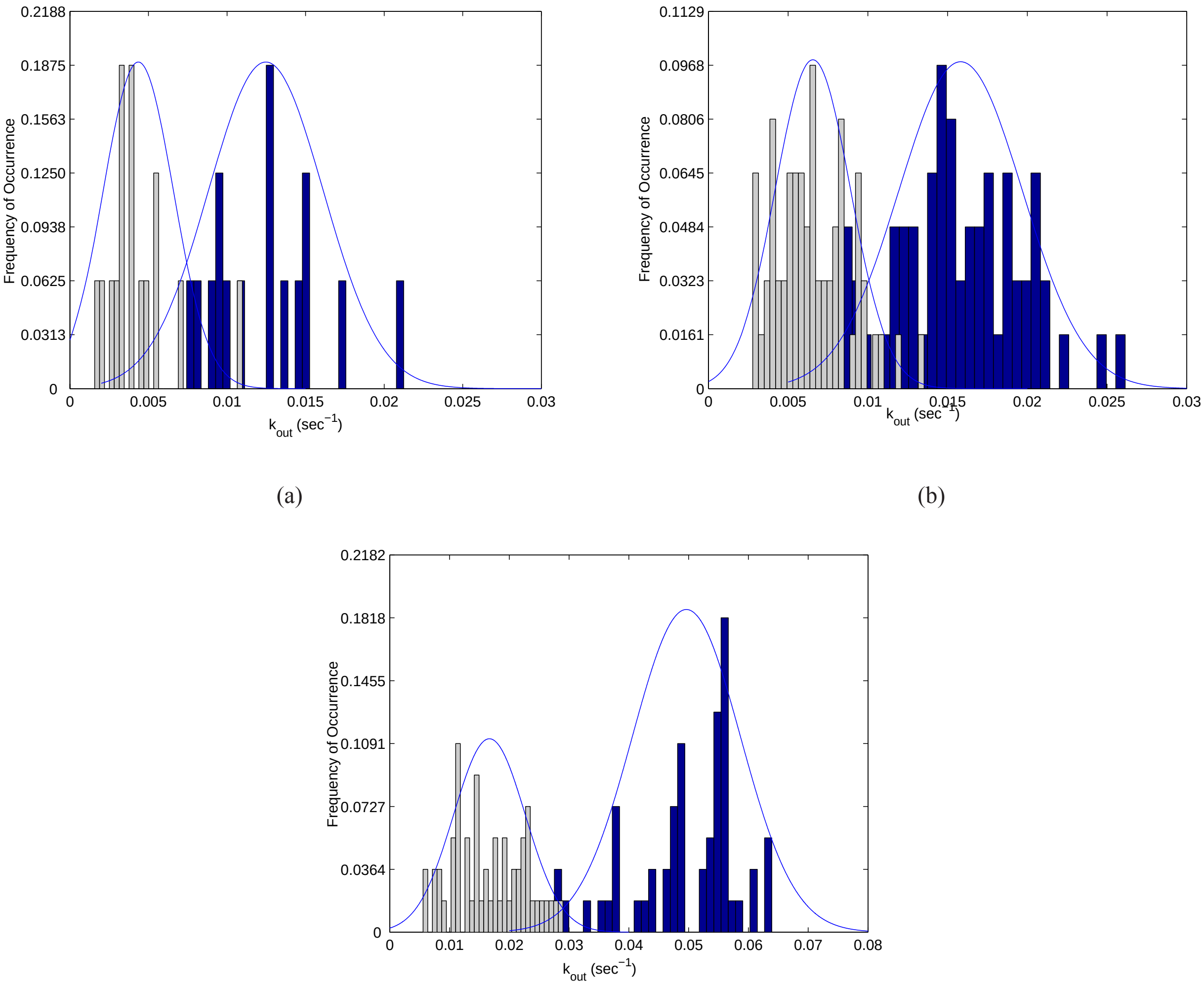


(a)

(b)

(c)

Fig. 18. The histograms of k_{in} for (a) Case 1, (b) Case 2, (c), Case 3 from inside (blue) and outside (gray) the tumor region (as indicated by circular/elliptical regions). The solid lines in figures shows the Gaussian fit.

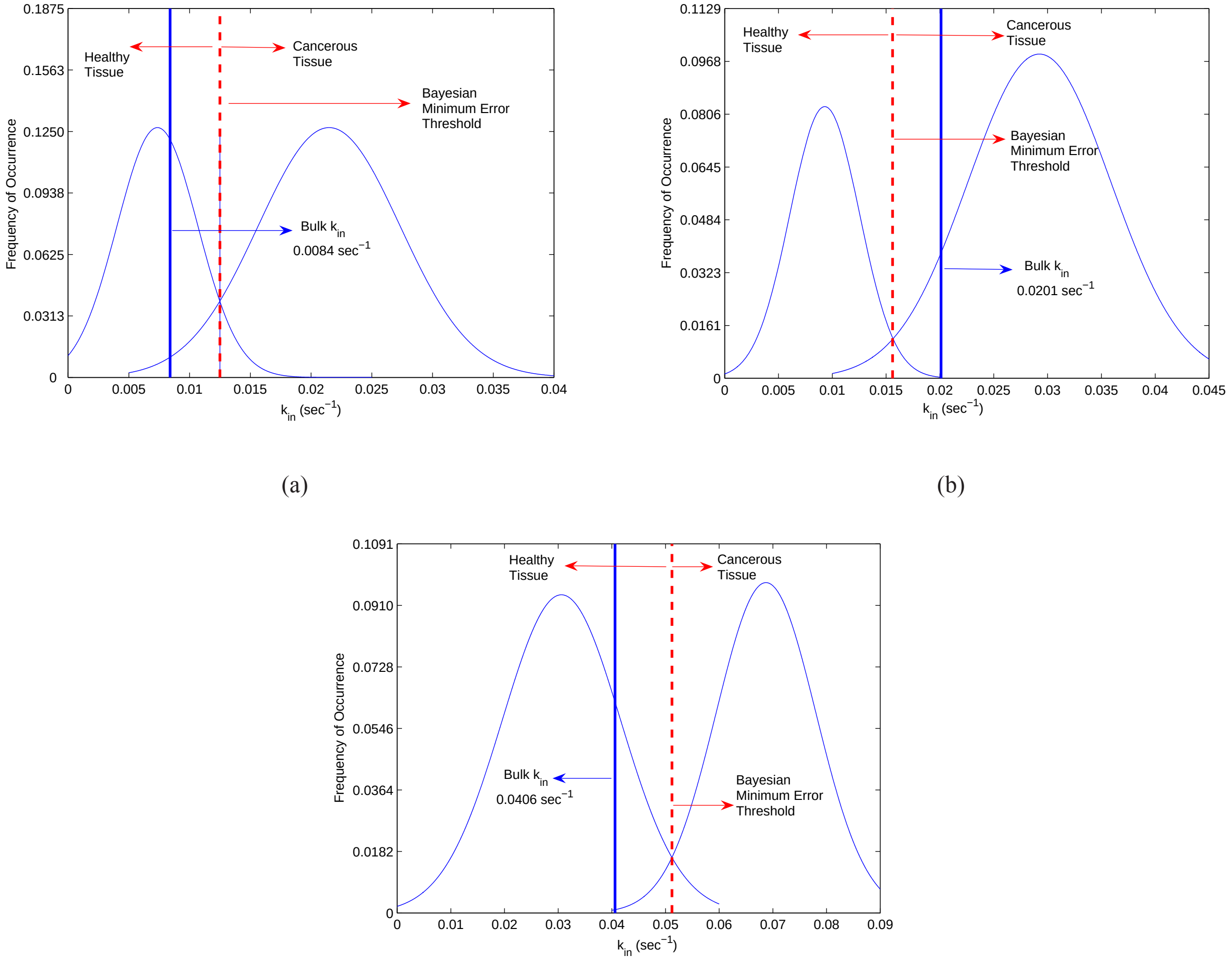


(a)

(b)

(c)

Fig. 19. The histograms of k_{out} for (a) Case 1, (b) Case 2, (c), Case 3 from inside (blue) and outside (gray) the tumor region (as indicated by circular/elliptical regions). The solid lines in figures shows the Gaussian fit.

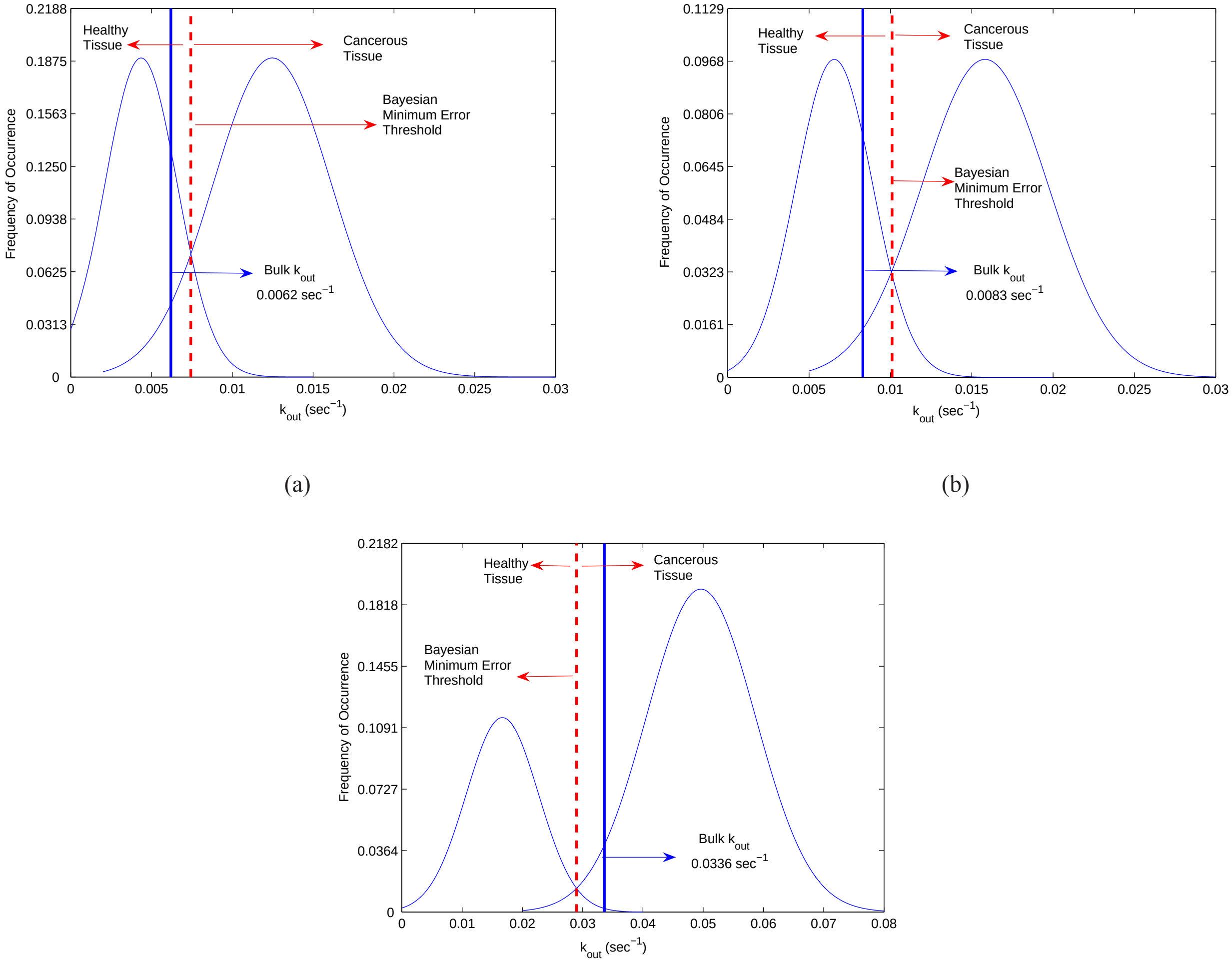


(a)

(b)

(c)

Fig. 20. Bulk k_{in} rates for (a) Case 1, (b) Case 2, (c), Case 3 together with histogram fits. The solid blue line indicates the bulk pharmacokinetic-rate, and the dashed red line indicates the Bayesian minimum error.



(a)

(b)

(c)

Fig. 21. Bulk k_{out} rates for (a) Case 1, (b) Case 2, (c), Case 3 together with histogram fits. The solid blue line indicates the bulk pharmacokinetic-rate, and the dashed red line indicates the Bayesian minimum error.

B. *AIM 3 - Tasks 1-4*

The SOW with regard to Aim 3 includes the following specific tasks:

- Task 1. Determine statistical variability of each NIR feature inside and outside the suspected tumor in an individual and evaluate the statistical significance of the measured difference with the instrumentation precision. *12-18th month*
- Task 2. Design statistical classifiers to determine the ROC of each NIR feature for an individual. *18-24th month*
- Task 3. Evaluate the ROC, positive predicted value (PPV) and negative predicted value (NPV) of various combinations of the NIR features for an individual. *24-27th month*
- Task 4. Investigate the significance of the measured difference between malignant and benign tumor patient groups for single and combined NIR features. *27-30th month*

Year 3 tasks involve Task 3 and 4.

This work describes the characterization efficiency of optical properties of breast tumors based on the features obtained using *in vivo* near-infrared (NIR) spectroscopy measurements. Three features, relative blood concentration, oxygen saturation and the size of the tumor, are used to diagnose benign and malignant tumors. The performance of the proposed set of features are evaluated by various classifiers using data acquired from 44 patients with malignant tumors, and 72 patients with benign tumors. The area under the receiver operating characteristics (ROC) curve of the scaled nearest mean classifier (NMSC) using the three features yields a value of 0.91 with a significance level of 0.05. Our results suggest that the features, relative blood concentration, and oxygen saturation can differentiate breast tumors with a relatively high precision.

1) Methods: Apparatus: In this study, a continuous wave (CW) near infrared spectrometer (NIRS) is used [**?**]. The apparatus includes a probe (Fig. 22). In the center of the probe there is a 3-wavelength light emitting diode (LED). The probe consisted of one multi-wavelength LED as a light source and 8 silicon diodes as detectors. The detectors surround the LED with a 4 cm radius. The light intensity from the detectors was adjusted to be approximately 1 volt and calibrated with a phantom with known absorption and scattering coefficients.

Patients and Protocol: This study includes two centers, namely, the Abramson Family Cancer Research Institute, Department of Radiology of the Hospital of University of Pennsylvania (HUP), and the Department of Gynecology of Leipzig University (DGLU). HUP provided 24 patients with malignant and 64 patients benign tumors. DGLU provided 20 patients with malignant and 6 benign tumors.

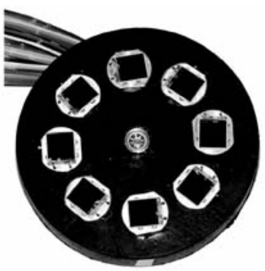


Fig. 22. The NIR probe with a multi-wavelength LED and 8 silicon diodes as detectors.

TABLE VI

AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F1-F2-F3: ΔBV, $\Delta Deoxy$, S

Type	**NMSC**	PAR	LDC	UDC	NEURC	QDC	NMC	KNNC
AUC	**0.9098**	0.9041	0.9017	0.8984	0.8864	0.8843	0.8807	0.8752

The measurements are taken on the breast with tumor. Then, the probe is transferred to the tumor free contralateral breast to include the mirror image location of the suspected cancer. The sensors giving the largest changes with respect to the mirror image position on the contralateral breast are related to the suspected cancer. *Feature Extraction*: In this study, three features, namely, relative blood concentration, ΔBV, oxygen saturation, $\Delta Deoxy$, and the size of the tumor, S, are used.

The features, ΔBV, and $\Delta Deoxy$ are obtained using

$$\Delta OD = \varepsilon \Delta CL \tag{14}$$

where OD is the optical density, ε is the extinction coefficient, C is blood concentration, L is the mean pathlength of photons, and Δ denotes relative change. Here, $\varepsilon \approx 1$ cm^{-1}, and $L = 4$ cm for a pathlength factor of 5.

Following (1), the relative blood concentration, ΔBV, and the oxygen saturation, $\Delta Deoxy$, can be approximated at two different wavelengths by

$$\Delta BV \propto 0.3\Delta OD_{730} + \Delta OD_{850} \tag{15}$$

$$\Delta Deoxy \propto 1.3\Delta OD_{730} + \Delta OD_{850} \tag{16}$$

where ΔOD_{730}, and ΔOD_{850} denote the relative changes in optical density at 730 nm and 850 nm, respectively.

ΔBV and $\Delta Deoxy$ can also be approximated by

$$\Delta BV \propto \Delta[Hb] + \Delta[HbO_2] \tag{17}$$

TABLE VII

AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F1-F2: ΔBV, $\Delta Deoxy$

Type	**NMSC**	PAR	LDC	UDC	NEURC	QDC	NMC	KNNC
AUC	**0.9001**	0.8993	0.8930	0.8908	0.8992	0.8821	0.8782	0.8645

$$\Delta Deoxy \propto \Delta[HbO_2] - \Delta[Hb] \tag{18}$$

where $\Delta[Hb]$, and $\Delta[HbO_2]$ denote the relative change in deoxyhemoglobin (Hb) and oxyhemoglobin ($Hb0_2$).

The concentrations of Hb, and HbO_2 in (17) and (18) are calculated by the Beer-Lambert Law given by

$$\Delta OD = log\frac{I_0}{I} \tag{19}$$

where I is light intensity after absorption and scattering, and I_0 is the baseline light intensity obtained from the contralateral breast, using known extinction coefficients of Hb, HbO_2 and differential pathlength factors [16].

Here, it is important to note that, ΔBV, and $\Delta Deoxy$ values are based on a lipid blood oxygen model. Thus the increments of BV and $Deoxy$ are relative to the contralateral breast:

$$\Delta BV = \Delta BV_{tumor} - \Delta BV_{contra} \tag{20}$$

$$\Delta Deoxy = \Delta Deoxy_{tumor} - \Delta Deoxy_{contra} \tag{21}$$

where ΔBV_{tumor}, ΔBV_{contra} are relative blood volume in the tumor breast and the mirror image position of the contralateral breast, respectively, and $\Delta Deoxy_{tumor}$, $\Delta Deoxy_{contra}$ are relative oxygen saturation in the tumor breast and the mirror image position of the contralateral breast, respectively.

Feature Analysis and Tumor Classification: In this subsection, we present the set of tumor classification features, and the malignancy differentiation criteria. F1 denotes ΔBV, F2 denotes $\Delta Deoxy$, and F3 denotes, S, size of the tumor. We evaluate the malignancy differentiation capability of the individual features and various combinations of the these features using a set of classifiers, namely, k-nearest neighbor classifier (KNNC), Parzen density based classifier (PAR), automatic neural network classifier (NEURC), normal densities based linear classifier (LDC), nearest mean classifier (NMC), scaled nearest mean classifier (NMSC), normal densities based quadratic classifier (QDC), uncorrelated normal densities based quadratic classifier (UDC). The more details information on these classifiers can be found in [17].

We evaluated the malignancy differentiation capability of the following individual and combined features:

F1: ΔBV

F2 : $\Delta Deoxy$

F3 : Tumor Size (S)

F1-F2: ΔBV and $\Delta Deoxy$

F1-F2-F3: ΔBV, $\Delta Deoxy$, and S

2) Statistical Analysis of Clinical Data: The evaluation is based on receiver operating characteristics (ROC) methodology. The ROC curve is obtained by plotting the probability of false positive rate versus the probability of detection. The evaluation of classification method is done using area under the ROC curve (AUC). First, we evaluated the classification performance of all three features. Table VI presents the AUC values for 8 different classifiers for all three features. The NMSC has the best performance in terms of classification with a AUC value of 0.9098 followed by the Parzen classifier with a AUC value of 0.9041.

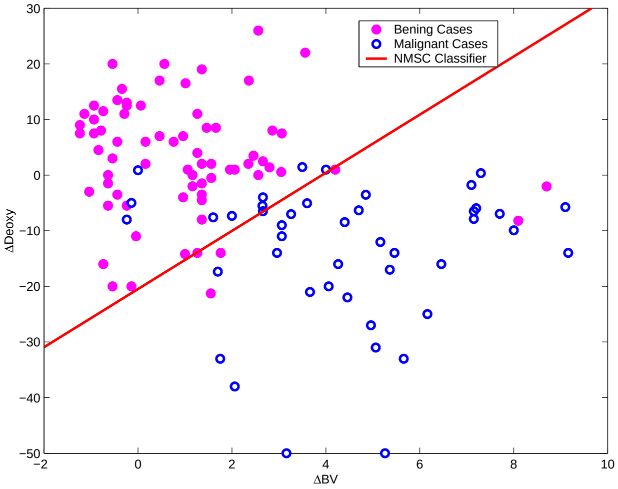


Fig. 23. Scaled Nearest Mean Classifier and F1-F2 2-D data clustering.

Next, we evaluated the performance of the two features measured by NIR spectroscopy. Table VII presents the AUC values for 8 different classifiers for features ΔBV and $\Delta Deoxy$. Again, the NMSC performed the best in terms of classification with a AUC value of 0.9001. Finally, we evaluated the individual classification performances of the three features. Table VIII presents the AUC values for 8 different classifiers for the feature ΔBV. The NMC has the best performance in terms of classification with a AUC value of 0.8832. Table IX presents the results 8 different classifiers for the feature $\Delta Deoxy$. The NMC has the best performance in terms of classification with a AUC value of 0.879. Table X presents the results 8 different classifiers for the feature S. The QDC has the best performance in terms of classification with a AUC value of 0.5612.

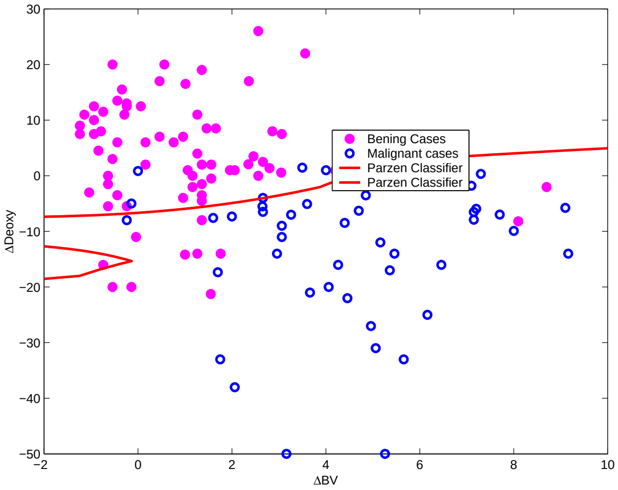


Fig. 24. Parzen and F1-F2 2-D data clustering.

TABLE VIII

AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F1: ΔBV

Type	NMSC	PAR	LDC	UDC	NEURC	QDC	**NMC**	KNNC
AUC	0.8817	0.8764	0.8807	0.8779	0.8513	0.8778	**0.8832**	0.8302

As it can be seen from Tables VI, and VII, the best performing feature set is the combination of the three features. We can also conclude from Table X that, the tumor size can not be used to differentiate healthy and diseased tissues with an AUC value of around 0.5. However, the combination set of optical indices, obtained using optical measurements, can differentiate breast tumors with a relatively high precision with a AUC value of 0.9. Similarly, optical indices, ΔBV and $\Delta Deoxy$, also performed well with AUC values of 0.883 and 0.879, respectively.

Figures 23, and 24 show the distribution of features ΔBV, and $\Delta Deoxy$ extracted from benign and malignant tumors. The thresholds were computed using scaled nearest mean classifier and Parzen classifiers. Figure 25 presents, the ROC curves for all three features, and the best two features, namely, $\Delta Deoxy$ and ΔBV. The observed area under the ROC curve for F1-F2-F3, and F1-F2 are 0.9098, and 0.9001, respectively. Figure 26 presents the ROC curves for individual features F1, and F2 using the nearest mean classifier . The observed area under the ROC curve for F1, and F2 are 0.8832 and 0.8790, respectively.

TABLE IX

AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F2: $\Delta Deoxy$

Type	NMSC	PAR	LDC	UDC	NEURC	QDC	**NMC**	KNNC
AUC	0.8787	0.8764	0.8776	0.8711	0.8491	0.8613	**0.8790**	0.8331

TABLE X

AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F3: S

Type	NMSC	PAR	LDC	UDC	NEURC	**QDC**	NMC	KNNC
AUC	0.5123	0.5292	0.4782	0.5429	0.5382	**0.5612**	0.5112	0.4827

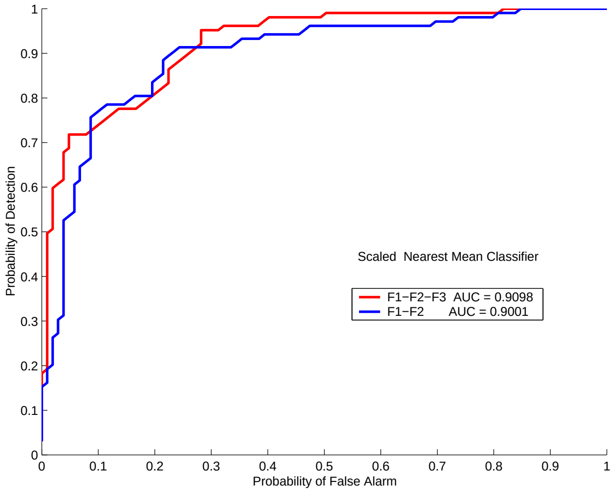


Fig. 25. ROC curves for F1-F2-F3 and F1-F2 using NMSC Classifier .

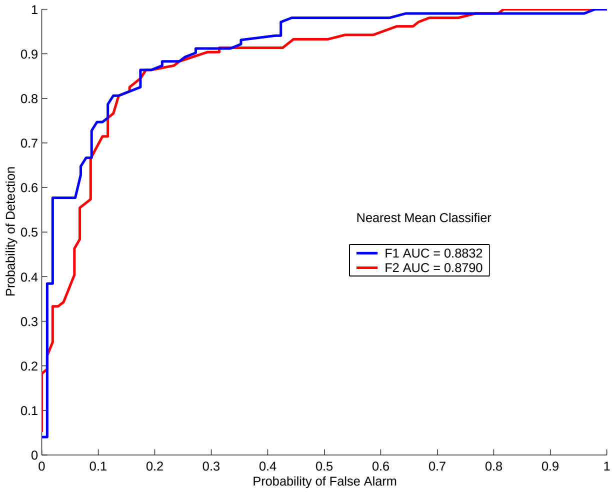


Fig. 26. ROC curves for F1and F2 using NMC Classifier .

III. KEY RESEARCH ACCOMPLISHMENTS

1) We developed a method of analyzing the discretization error in diffuse optical image reconstruction and developed a novel adaptive mesh generation algorithm based on MR a priori anatomical images. We applied our method to diffuse optical absorption imaging. The results of this work is published in two journal articles in Inverse Problems (8),(9).
2) We developed an adaptive mesh generation algorithms for diffuse optical imaging for simultaneous reconstruction of diffusion and absorption images. This work is currently being prepared as a journal paper.
3) We demonstrated, for the first time in the literature, the value of spatially resolved pharmacokinetic-rates as opposed to bulk-rates using in vivo breast cancer patient data. This work is submitted as a journal paper to IEEE Transactions in Biomedical Engineering (1).
4) We developed methods of reconstructing pharmacokinetic-rate images directly from NIR boundary measurements and applied our technique to in vivo breast cancer data. The results show that the technique is robust and provides better signal-to-noise ratio images. We are in the process of assessing the implications of this results in breast cancer diagnosis and staging.
5) We further analyze the NIR parameters collected from 116 patients using NIR spectroscopy and reported a short version of our results in (2). This work is currently being prepared as a journal paper.

IV. REPORTABLE OUTCOMES

Complete list of outcomes is given below:

1) B. Alacam, B. Yazici, X. Intes, B. Chance, S. Nioka, “Pharmacokinetic-Rate Images of Indocyanine Green for Breast Tumors using Near Infrared Optical Methods,” in review *IEEE Transaction in Biomedical Engineering*, May 2007.
2) B. Alacam, B. Yazici. B. Chance, S. Nioka, “Characterization of Breast Tumors with NIR Methods using Optical Indices,” to be published in *29th IEEE EMBS Conference*, 2007.
3) B. Alacam, B. Yazici X. Intes, B. Chance, “Extended Kalman filtering for the modeling and analysis of ICG pharmacokinetics in cancerous tumors using NIR optical methods,” *IEEE Transactions in Biomedical Engineering*, Vol. 53-10, pp: 1861-1871, 2006.
4) B. Alacam, B. Yazici, X. Intes, S. Nioka, B. Chance, “Spatially resolved pharmacokinetic rate images of ICG

using Near Infrared optical methods", Proc. of 2006 *SPIE Photonic West*, San Jose, California USA, 21- 26 January 2006, vol. 6088. pp. 455-464.

5) B. Alacam, B. Yazici, X. Intes, S. Nioka, B. Chance, "Direct reconstruction of pharmacokinetic rate Images of Indocyanine Green in fluorescence molecular tomography" *Proceedings of Biomedical Optics Topical Meeting*, Fort Lauderdale, Florida, USA , 19-22 March, 2006, No. SH65.
6) B. Alacam, B. Yazici, A. Serdaroglu, X. Intes, B. Chance, "Reconstruction of spatially resolved pharmacokinetic rate images of fluorescence agents in FDOT" *Proc. of EMBS-28th* Anniversary Conference, September, 2006, New York, pp. 5627–5630.
7) B. Alacam, B. Yazici, X. Intes, B. Chance, "Direct reconstruction of spatially resolved pharmacokinetic rate images of fluorescence agents" to be submitted to *IEEE Transactions in Medical Imaging*.
8) M. Guven, B. Yazici, K. Kwon, E. Giladi, X. Intes, Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: Part II," Inverse Problems, Vol. 23, pp: 1135-1160, May 2007.
9) M. Guven, B. Yazici, K. Kwon, E. Giladi, X. Intes, Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: Part I," Inverse Problems, Vol. 23, pp: 1115-1133, May 2007.
10) K. Kwon, B. Yazici, and M. Guven Two-level domain decomposition methods for diffuse optical tomography, Inverse Problems, Vol. 22, pp: 1533-1559, 2006.
11) K. Kwon and B. Yazici, Born expansion and Frechet derivatives in diffuse optical tomography, submitted to Inverse Problems, May 2006.
12) M. Guven, B. Yazici, K. Kwon, E. Giladi, and X. Intes, Adaptive mesh generation for DOT to reduce the error resulting from discretization", IEEE Symposium in Biomedical Imaging, April 2007. invited.
13) M. Guven, B. Yazici and V. Ntziachristos, Fluorescence diffuse optical Image reconstruction with a priori information, Proceedings of SPIE Photonics West 2007. Invited.

V. CONCLUSIONS

In the last 12 months, we continued to analyze the 116 breast cancer patient data obtained using NIR spectroscopy and refine our imaging algorithms. We showed that excluding the pharmacokinetic rate images, the NIR parameters alone can provide an AUC value of 0.9 or higher. Additionally, we analyzed the three in vivo breast cancer patient data acquired using a dynamic time-resolved tomographic system. The patients were injected ICG (an optical contrast agent) and diffuse optical tomographic data was collected before and after the injection. We reconstructed the spatially resolved pharmacokinetic-rates of ICG using variety of techniques and analyzed the resulting images in terms of diagnostic value. Our analysis shows that the spatially-resolved pharmacokinetic rates provide superior diagnostic information in terms of specificity and sensitivity as compared to a single set of pharmacokinetic-rates obtained using spectroscopic techniques.

We continued to refine our NIR imaging algorithms. We published a novel adaptive meshing algorithms for diffuse optical absorption imaging using the MR *a priori* information. We extended our algorithms for simultaneous reconstruction of diffusion and absorption images. We also developed methods of pharmacokinetic-rate images directly from NIR boundary measurements which provide robust reconstruction with higher SNR than the pharmacokinetic-rate images obtained from absorption images. We are currently assessing the value of this technique in breast cancer diagnosis in terms of sensitivity and specificity.

REFERENCES

[1] Folkman, J., Tumor angiogenesis: therapeutic implications. N. Engl. J. Med., 1971. 285(21): p. 1182-6.

[2] Folkman, J., What is the evidence that tumors are angiogenesis dependent? J. Natl. Cancer inst., 1990. 82(1): p. 4-6.

[3] Blood, C. H. and B. R. Zetter, Tumor interactions with the vasculature: angiogenesis and tumor metastasis. Biochem. Biophys. Acta, 1990. 1032(1): p. 89-118.

[4] Vaupel, P., et al., Current status of knowledge and critical issues in tumor oxygenation. Results from 25 years research in tumor pathophysiology. Adv. Exp. Med. Biol., 1998. 454: p. 591-602.

[5] Vaupel, P., et al., Oxygenation of human tumors: evaluation of tissue oxygen distribution in breast cancers by computerized O2 tension measurements. Cancer Res., 1991. 51(12): p. 3316-22.

[6] Runkel, S., et al., Oxygenations of mammary tumors as evaluated by ultrasound-guided computerized-pO2-histograpy. Adv. Exp. Med. Biol., 1994. 345: p. 451-8.

[7] Pogue, B. W., et al., Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: pilot results in the breast. Radiology, 2001. 218(1): 261–6.

[8] Gurfinkel M et al 2000 Pharmokinetics of ICG and HPPH-car for the detection of normal and tumour tissue using fluorescence, near-infrared reflectance imaging: a case study Photochem. Photobiol. 72 94-102

[9] Harms S E. Breast magnetic resonance imaging. Seminars in Ultrasound, CT & MR 1998; 19:104-120.

[10] Harms SE, Flamig DP, Hesley KL, et al. MR imaging of the breast with rotating delivery of excitation off resonance: clinical experience with pathologic correlation. Radiology 1993; 187:493-501.

[11] Williams MB, Pisano ED, Schnall MD, Fajardo LL. Future directions in imaging of breast diseases. Radiology 1998; 206:297-300.

[12] Swift M, Morrell D, Massey RB, Chase CL. Incidence of cancer in 161 families affected by ataxia-telangiectasia. New England Journal of Medicine 1991; 325:1831-1836.

[13] Esserman L, Hylton N, Yassa L, et al. Utility of magnetic resonance imaging in the management of breast cancer: Evidence for improved preoperative staging. J Clin Oncol 1999; 17:110-119.

[14] Robinson, K. J., C. J. Kotre and K. Faulkner. ”The use of contrast-detail test objects in the optimization of optical density in mammography.” Br. J. Radiol. 68: 277-282, 1995.

[15] Moran, G. R. and F. S. Prato. ”Modeling tissue contrast agent concentration: a solution to the tissue homogeneity model using a simulated arterial input function.” Mag. Res. Med. 45:42-45, 2001.

[16] Fantini S, Hueber D, Franceschini MA, Gratton E, Rosenfeld W, Stubblefield PG, Maulik D, Stankovic MR., ”Non-invasive optical monitoring of the newborn piglet brain using continuous-wave and frequency-domain spectroscopy,” *Phys. Med. Biol.*, Vol. 44, pp. 1543-1563, 1999.

[17] Duda RO, Hart PE, Stork DG. Pattern Classification. New York: Wiley-Interscience, 2000.

[18] X. Intes, J. Ripoll, Y. Chen, S. Nioka, A. G. Yodh, B. Chance, “*In vivo* continuous-wave optical breast imaging enhanced with Indocyanine Green,” *Med. Phys.* vol. 30-6, pp. 1039-1047, 2003.

[19] B. Alacam, B. Yazici, X. Intes, B. Chance, “Extended Kalman Filtering for the Modeling and Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods”, *Transactions in IEEE Biomedical Engineering*, vol. 53, no. 10, pp. 1861-1871, 2006.

[20] M. L. J. Landsman, G. Kwant, G. A. Mook, W. G. Zijlstra,“Light- absorbing properties, stability, and spectral stabilization of indocyanine green,” *J. Appl. Physiol.*, vol. 40, pp. 575-583, 1976.

[21] M. O’Leary,‘Imaging with Diffuse Photon Density Waves,” in PhD Thesis, Dept. Physics & Astronomy, U. of Pennsylvania, May 1996.

[22] M. Gurfinkel, A. B. Thompson, W. Ralston, T. L. Troy, A. L. Moore, T. A. Moore, J. D. Gust, D. Tatman, J. S. Reynolds, B. Muggenburg, K. Nikula, R. Pandey, R. H. Mayer, D. J. Hawrysz, and E. M. Sevick-Muraca, “Pharmacokinetics of ICG and HPPH-car for the detection of normal and tumor tissue using fluorescence, near-infrared reflectance imaging: a case study,” *Photochem. Photobiol.*, Vol. 72, pp. 94-102, 2000.

[23] V. Ntziachristos, B. Chance, and A. Yodh,“Differential diffuse optical tomography,” *Opt. Express*, vol. 5, pp. 230-242, 1999.

[24] J. Ripoll, V. Ntziachristos, R. Carminati, and M. Nieto-Vesperinas, “Kirchhoff approximation for diffusive waves,” *Phys. Rev.* vol. 64, pp. 051917, 2001.

[25] J. Ripoll, V. Ntziachristos, J. P. Culver, D. N. Pattanayak, A. G. Yodh, and M. Nieto-Vesperinas, “Recovery of optical parameters in multiplelayered diffusive media: theory and experiments,” *J. Opt. Soc. Am.* vol. A 18, pp. 821830, 2001.

[26] D. Boas, T. Gaudette, and S. Arridge, “Simultaneous imaging and optode calibration with diffuse optical tomography,” *Opt. Express* vol. 8, pp. 263270, 2001.

[27] P. Hansen and D. OLeary, “The use of the L-curve in the regularization of discrete ill-posed problems,” *SIAM J. Sci. Comput.*, vol. 14, pp. 1487 1503, 1993.

[28] C. Chen, *Linear System Theory and Design*, Oxford University Press, New York, 1999.

[29] L. Ljung, “Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems,” *IEEE Tran. Automa. Control*, vol. AC-24, no. 1, pp. 36-50, 1979.

[30] R. Togneri, and L. Deng, “Joint State and Parameter Estimation for a Target-Directed Nonlinear Dynamic System Model,” *IEEE Tran. on Sig. Proc.*, vol. 51, no. 12, pp. 3061-3070, 2003.

[31] L. Nelson, and E. Stear, “The Simultaneous On-Line Estimation of Parameters and States in Linear Systems,” *IEEE Tran. on Auto. Control*, vol.21, pp. 94-98, 1976.

[32] S. Nioka, Y. Yung, M. Schnall, S. Zhao, S. Orel, C. Xie, B. Chance, and S. Solin,“Optical imaging of breast tumor by means of continuous waves,” *Adv. Exp. Med. Biol.*, vol. 411, pp. 227-232, 1997.

[33] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, 1990.

Appendix
Key Publications

Pharmacokinetic-Rate Images of Indocyanine Green for Breast Tumors using Near Infrared Optical Methods

Burak Alacam[1], *Student Member, IEEE*, *Birsen Yazıcı[1,2], *Senior Member, IEEE*
Xavier Intes[2], Britton Chance[3], and Shoko Nioka[3]

[1]Department of Electrical, Computer and System Engineering,
Rensselaer Polytechnic Institute 110 8th Street, Troy, NY 12180
Phone: (518) 276-2905, Fax: (518) 276-6261
E-mail: yazici@ecse.rpi.edu
[2]Department of Biomedical Engineering
Rensselaer Polytechnic Institute, Troy, NY
[3]Department of Biochemistry and Biophysics,
University of Pennsylvania, Philadelphia, PA

This work was supported by U.S. Army Medical Research Acquisition Activity under grant W81XWH-04-1-0559, and the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation, Award Number EEC-9986821. *Corresponding author

Abstract

In this paper, we develop a method of forming pharmacokinetic-rate images of indocyanine green (ICG) and apply our method to three sets of *in vivo* data obtained from breast cancer patients. To form pharmacokinetic-rate images, we first obtain a sequence of ICG concentration images using the differential diffuse optical tomography technique. We next employ a two-compartment model composed of plasma, and extracellular-extravascular space (EES), and estimate the pharmacokinetic-rates and concentrations in each compartment using the extended Kalman filtering framework. The pharmacokinetic-rate images of the three patient show that the rates from inside and outside the tumor region are statistically different with very low probability of false negative and false positive. Additionally, the ICG concentrations in plasma, and the EES compartments are higher around the tumor region tumor agreeing with the hypothesis that around the tumor region ICG may act as a diffusible extravascular flow in compromised capillary of cancer vessels. Our study shows that the pharmacokinetic-rate images provide superior information than single set of pharmacokinetic-rates estimated from the entire breast tissue for breast cancer diagnosis.

Keywords: Pharmacokinetic analysis, indocyanine green, differential diffuse optical tomography, breast cancer.

I. INTRODUCTION

Near infrared (NIR) diffuse optical imaging offers several advantages over other imaging modalities [1]–[7]. NIR techniques are minimally invasive, easy to use, relatively inexpensive, and can be made portable. Moreover, optical techniques, when coupled with contrast agents, have the potential to provide molecular/cellular level information, which can improve cancer detection, staging, and treatment monitoring [3], [4], [8]–[10].

Among many commercially available optical contrast agents, only indocyanine green (ICG) is approved for use in humans by the Food and Drug Administration [11]–[13]. ICG is a blood pooling agent that has different delivery behavior between normal and cancer vasculature. In normal tissue, ICG acts as a blood flow indicator in tight capillaries of normal vessel. However in cancerous tumor, ICG may act as a diffusible (extravascular) flow in leaky capillary of cancer vessels [8], [10], [14]. Therefore, pharmacokinetics of ICG has the potential to provide new tools for tumor detection, diagnosis, and staging.

One approach to analyze pharmacokinetics of contrast agents is the compartmental modeling [15]–[17]. A number of studies using compartmental modeling were reported to show the feasibility of ICG pharmacokinetics in tumor characterization [8]–[10]. Cuccia et al. [8] presented a study of the dynamics of ICG in an adenocarcinoma rat tumor model using a two-compartment model. Intes et al. [9] presented the uptake of ICG in breast tumors using a continuous wave diffuse optical tomography apparatus using a two-compartment model. We recently introduced the extended Kalman filtering (EKF) framework to model and estimate the ICG pharmacokinetics and tested three different compartmental models for the ICG pharmacokinetics using the *in vivo* NIR data collected from Fischer rats with cancerous tumors [10]. Our study suggests that the pharmacokinetic-rates out of the vasculature are higher in edematous tumors as compared to necrotic tumors.

In all the studies described above, the pharmacokinetic-rates are assumed to be constant over a tissue volume that may be as large as the entire imaging domain. However, pharmacokinetic-rates are expected to be different in healthy and cancerous tissue as reported in positron emission tomography (PET), and magnetic resonance imaging (MRI) literature. In [18]–[20], it was shown that the spatially resolved pharmacokinetic-rates increase sensitivity and specificity for breast cancer diagnosis. For example, in [18], Sun et al. showed that FAU (1-2’-deoxy-2’-fluoro- β-D-arabinfuranosyl urasil, a PET contrast agent) accumulation in tumor region is significantly higher when compared to normal breast tissue based on pharmacokinetic-rate images. Mussurakis et al. [19] showed that the pharmacokinetics of gadolinium-DTPA (an MRI contrast agent) can be used to differentiate between malignant and benign breast tumors with a high accuracy. It has also been shown that the spatially resolved image interpretation is superior to the isolated use of quantitative pharmacokinetic-rates. Since gadolinium-DTPA kinetics is expected to be similar to ICG, [19], and [20] suggest that the spatially resolved ICG pharmacokinetic analysis may provide better sensitivity and specificity.

In the area of diffuse NIR spectroscopy and imaging, a number of studies on spatially resolved pharmacokinetic-rates has been reported [21], [22]. Gurfinkel et al. [21] presented *in vivo* NIR reflectance images of ICG pharmacokinetics to discriminate canine adenocarcinoma (located at 0.5-1 cm depth) from normal mammary tissue. These images were generated by a non-tomographic technique using a CCD camera that is suitable only to image tumors close to surface. In [22], Milstein et al. presented a Bayesian tomographic image reconstruction method to form pharmacokinetic-rate images of ICG based on fluorescence diffuse optical tomography. Numerical simulations show that the method provides good contrast. However, no real data experiments were presented to study the clinical value of spatially resolved pharmacokinetic-rates.

In this paper, we present a method of forming pharmacokinetic-rate images and report spatially resolved pharmacokinetic-rates of ICG using the *in vivo* NIR data acquired from three breast cancer patients. To the best of our knowledge, our work is the first presenting the pharmacokinetic-rate images of an optical contrast agent using *in vivo* breast data based on tomographic techniques. We first develop a set of spatio-temporally resolved ICG concentration images based on differential diffuse optical tomography. We model the ICG pharmacokinetics by a two-compartment model composed of plasma and extracellular-extravascular space (EES) compartments. We then estimate the ICG pharmacokinetic-rates and the concentrations in different compartments based on the EKF framework [10]. We show that the pharmacokinetic-rates from inside and outside the tumor region are statistically different. Total error values using the Bayesian minimum error classifier show that spatially resolved pharmacokinetic-rate analysis provides very good sensitivity and specificity. We also estimate a single set of pharmacokinetic-rates (bulk pharmacokinetic-rates) for the entire breast tissue. Our study indicates that spatially resolved pharmacokinetic-rates provides more consistent and superior diagnostic information as compared to the bulk pharmacokinetic-rates.

The rest of the paper is organized as follows: In Section II, we present the reconstruction of ICG concentration images. In Section III, we present modeling and estimation of ICG pharmacokinetics using the EKF framework. In Section IV, we describe the clinical results obtained from breast cancer patients. Section V summarizes our results and conclusion.

II. RECONSTRUCTION OF BULK ICG CONCENTRATION IMAGES

In our data collection process, a sequence of boundary measurements are collected over a period of time. Each set of measurements are used to form a frame of the ICG concentration images. The resulting sequence of ICG concentration images are then used to form pharmacokinetic-rate images. To reconstruct each frame of the ICG concentration images, we follow a static reconstruction approach and use differential diffuse optical tomography (DDOT) technique [9], [23].

In DDOT, two sets of excitation measurements are collected corresponding to before and after the ICG injection, and the ICG concentration is determined by the perturbation method [9], [23]. The photon propagation before and after the injection is modeled by the following diffusion equation:

$$\nabla \cdot [D_x \nabla \Phi_x^{\pm}(r,\omega) - [\mu_{ax}^{\pm}(r) + j\omega/c]\Phi_x^{\pm}(r,\omega) = 0, \qquad r \in \Omega \subseteq \mathbb{R}^3 \tag{1}$$

with Robin-type boundary conditions:

$$2D_x \frac{\partial \Phi_x^{\pm}}{\partial \nu} + \rho \Phi_x^{\pm} = -S(r,\omega), \qquad r \in \partial\Omega. \tag{2}$$

where x stands for the excitation, c is the speed of light inside the medium Ω; ω denotes the modulation frequency of the source, $\mu_{ax}^{-}(r)$ and $\mu_{ax}^{+}(r)$ are the absorption coefficients before and after the ICG injection, D_x is the diffusion coefficient which is assumed to be constant, $\Phi_x^{\pm}(r,\omega)$ optical field at location r before and after the ICG injection. Here, ν denotes the outward normal to the boundary $\partial\Omega$ of Ω, ρ is a constant representing the refractive index mismatch between the two regions separated by $\partial\Omega$, and $S(r,\omega)$ is the excitation source on the boundary.

The absorption coefficient after the injection μ_{ax}^{+} are modeled as a sum of the absorption coefficient of the medium before the ICG injection μ_{ax}^{-} and the perturbation caused by the ICG $\Delta\mu_{ax}(r)$:

$$\Delta\mu_{ax}(r) = \mu_{ax}^{+}(r) - \mu_{ax}^{-}(r), \qquad r \in \Omega \subseteq \mathbb{R}^3. \tag{3}$$

In the forward model, the analytical solutions of the heterogonous diffusion equation given in (1) is derived using first order Rytov approximation [9]. The sample volume is divided into a set of voxels and the measurements are related to the relative absorption coefficients of each voxel by a system of linear equations. The shape of the breast was approximated as a cylinder and the Kirchhoff approximation [24], [25] for diffuse waves was used to model the interaction of light with boundaries. In order to account for the biological noise, the forward model was implemented with coupling coefficient technique [26].

Here, the Rytov-type measurements, which are defined by the natural logarithm of the ratio of the post-ICG measurements to the pre-ICG measurements were used [23]. Let $\Psi_x(r_d, \omega; r_s)$ denote the Rytov-type measurements at location r_d due to source at r_s. The linearized relationship between the differential absorption coefficient and measurements is given by [27],

$$\Psi_x(r_d, \omega; r_s) = -\frac{1}{\Phi_x^-(r_d, \omega; r_s)} \int_\Omega G_x^-(r - r_d, \omega; r_s) I_x(r) \Phi_x^-(r, \omega; r_s) d^3r \tag{4}$$

where $\Phi_x^-(r, \omega; r_s)$ is the photon density obtained at the excitation wavelength before ICG injection, $I_x(r) = c\Delta\mu_{ax}(r)/D_x$, and $G_x^-(r - r_d, \omega; r_s)$ is the Green's function of (1) for a source at r_s before the injection describing the propagation of light from the heterogeneity r to the detector at r_d.

We address the inverse problem of recovering $\Delta\mu_{ax}$ from Rytov-measurements Ψ_x based on the forward model (4) using the singular value decomposition of the Moore-Penrose generalized system. We use a zeroth-order Tikhonov regularization to stabilize the inversion procedure. We apply the L-curve method to an experimental model reconstruction and derive the best regularization parameter using a curvature function as described in [28]. A detailed discussion of the forward and inverse models used for the reconstruction of differential absorption coefficients ($\Delta\mu_{ax}$) can be found in [9].

To construct a set of ICG concentration images, we use the linear relationship between the differential absorption coefficients and ICG concentrations [29]:

$$\Delta\mu_a(r) = \ln 10 \; \epsilon_\lambda \; m(r) \; = 2.303 \; \epsilon_\lambda \; m(r) \tag{5}$$

where ϵ_λ is the extinction coefficient of ICG at the wavelength 805nm, $m(r)$ is the bulk ICG concentration in the tissue, and $\Delta\mu_a(r)$ is as defined in (3).

III. MODELING AND ESTIMATION OF ICG PHARMACOKINETICS

A. Two-compartment Model of ICG Pharmacokinetics

Using the method outlined in Section 2, we reconstruct a sequence of ICG concentration images. As an example, Figure 3 shows a set of images reconstructed from *in vivo* breast data.

Our objective is to model the pharmacokinetics of ICG at each voxel of ICG concentration images using compartmental modeling. To do so, we first extracted the time varying ICG concentration curves for each voxel from the sequence of ICG concentration images. An example of such a curve is shown in Figure 6. We next fit a two-compartment model to each ICG concentration curve [10], [21]. Figure 1 shows the two-compartment model for ICG kinetics. Here, plasma and extracellular-extravascular space (EES) constitute the two compartments. C_p and C_e represent the ICG concentrations in plasma and the EES, respectively. The rates k_{in} and k_{out} govern the leakage into and the drainage out of the EES. The parameter k_{elm} describes the ICG elimination from the body through kidneys and livers.

Using the two-compartment model introduced in [10], ICG transition between plasma and the EES can be modeled as follows:

$$\begin{bmatrix} \dot{C}_e(t) \\ \dot{C}_p(t) \end{bmatrix} = \begin{bmatrix} -k_{out} & k_{in} \\ k_{out} & -(k_{in}+k_{elm}) \end{bmatrix} \begin{bmatrix} C_e(t) \\ C_p(t) \end{bmatrix} + \boldsymbol{\omega}(t), \tag{6}$$

where $C_p(t)$ and $C_e(t)$ represent the ICG concentrations in plasma and the EES at time $t \in [T_0, T_1]$, respectively. The rates k_{in}, k_{out}, and k_{elm} have a unit of sec^{-1}. They are defined as the permeability surface area products given by $PS\rho$, where P is the capillary permeability constant, S is the capillary surface area, and ρ is the tissue density. Here, $\boldsymbol{\omega}(t)$ is uncorrelated zero-mean Gaussian process with covariance matrix **Q** representing the model mismatch.

The actual total ICG concentration in the tissue is a linear combination of plasma and the EES ICG concentrations modeled as:

$$m(t) = \begin{bmatrix} v_e & v_p \end{bmatrix} \begin{bmatrix} C_e(t) \\ C_p(t) \end{bmatrix} + \eta(t) \tag{7}$$

where $m(t)$, $C_e(t)$, and $C_p(t)$ are defined in (5) and (6); v_p and v_e are plasma and the EES volume fractions, respectively; and $\eta(t)$ is uncorrelated zero-mean Gaussian process with covariance matrix **R**, representing the measurement noise.

B. Estimation of ICG Pharmacokinetics using Extended Kalman Filtering

In matrix-vector notation, (6) and (7) can be expressed as:

$$\dot{\mathbf{C}}(t) = \mathbf{K}(\boldsymbol{\alpha})\mathbf{C}(t) + \boldsymbol{\omega}(t), \tag{8}$$

$$m(t) = \mathbf{V}(\boldsymbol{\alpha})\mathbf{C}(t) + \eta(t),$$

where $\mathbf{C}(t)$ denotes the concentration vector with elements $C_e(t)$, and $C_p(t)$; $\mathbf{K}(\boldsymbol{\alpha})$ is the system matrix, $\mathbf{V}(\boldsymbol{\alpha})$ is the measurement matrix as defined in equation (7), and $\boldsymbol{\alpha}$ is the parameter vector given by

$$\boldsymbol{\alpha} = [k_{out} \;\; k_{in} \;\; k_{elm} \;\; v_e \;\; v_p]^T. \tag{9}$$

The ICG measurements in (8) are collected at discrete time instances, $t = kT$, $k = 0, 1, ...$, where T is the sampling period. Therefore, the continuous model described in (8) is discretized. We can express the discrete compartmental model as follows:

$$\mathbf{C}_d(k+1) = \mathbf{K}_d(\boldsymbol{\theta})\mathbf{C}_d(k) + \boldsymbol{\omega}_d(k), \tag{10}$$

$$\mathbf{m}(k) = \mathbf{V}_d(\boldsymbol{\theta})\mathbf{C}_d(k) + \boldsymbol{\eta}_d(k),$$

where $\mathbf{K}_d(\boldsymbol{\theta}) = e^{\mathbf{K}(\boldsymbol{\alpha})}$ is the discrete time system matrix; $\mathbf{V}_d(\theta) = \mathbf{V}(\boldsymbol{\alpha})$ is the discrete measurement matrix; $\boldsymbol{\omega}_d(k)$ and $\boldsymbol{\eta}_d(k)$ are zero-mean Gaussian white noise processes with covariances matrix $\mathbf{Q}_d$ and variance $\mathbf{R}_d$, respectively. The vector $\boldsymbol{\theta}$ is composed of parameters τ_{ij} which are functions the pharmacokinetic-rates and volume fractions:

$$\boldsymbol{\theta} = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{21} & \tau_{22} & v_e & v_p \end{bmatrix}^T. \tag{11}$$

We first estimate τ_{ij}'s, $i, j = 1, 2$ and then compute the pharmacokinetic-rates k_{in}, k_{out} and k_{elm} [10], [30]. The explicit form of the discrete state-space model is given as follows:

$$\begin{bmatrix} C_e(k+1) \\ C_p(k+1) \end{bmatrix} = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{bmatrix} \begin{bmatrix} C_e(k) \\ C_p(k) \end{bmatrix} + \boldsymbol{\omega}_d(k) \tag{12}$$

$$m(k) = \begin{bmatrix} v_e & v_p \end{bmatrix} \begin{bmatrix} C_e(k) \\ C_p(k) \end{bmatrix} + \eta_d(k).$$

We estimate the parameter vector $\boldsymbol{\theta}$ and concentration vector $\mathbf{C}_d$ by using the EKF framework. The EKF is a recursive modeling and estimation method with numerous advantages in ICG pharmacokinetic modeling [10]. These include effective modeling of multiple compartments, and multiple measurement systems in the presence of measurement noise and uncertainties in the compartmental model dynamics, simultaneous estimation of model parameters and ICG concentrations in each compartment, statistical validation of estimated concentrations and error bounds on the model parameter estimates, and incorporation of available a priori information about the initial conditions of the permeability rates into the estimation procedure.

When both states (ICG concentrations) and model parameters (pharmacokinetic-rates and volume fractions) are unknown, a linear state-space model can be regarded as a non-linear model; the linear system parameters and states combine to form the new states of the non-linear model. This system is then linearized and the new unknown states are found using the EKF estimator [10], [31]–[33]. In EKF framework, $\boldsymbol{\theta}$ can be treated as a random process with the following model:

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \varsigma_d(k), \tag{13}$$

where $\varsigma_d(k)$ is a zero-mean Gaussian process with covariance matrix $\mathbf{S}_d$.

Table 1 summarizes the joint estimation of pharmacokinetic-rates and ICG concentration in different compartments. In Table 1, $\hat{\mathbf{C}}_d(k|k-1)$ is the state estimate propagation at step k given all the measurements up to step $k-1$; $\hat{\mathbf{C}}_d(k)$ is the state estimate update at step k; $\mathbf{P}_{k,k-1}$ denotes the error covariance propagation at step k given all the measurements up to step $k-1$; $\mathbf{P}_{k,k}$ is the error covariance update at step k; $\mathbf{S}_d$ is the preassigned covariance matrix of $\varsigma_d(k)$; $\mathbf{J}_k$ is the Jacobian matrix due to iterative linearization of the state equation at step k; $\mathbf{G}_k$ is the recursive Kalman gain at step k; $\mathbf{R}_d$ is the covariance matrix of the measurements; $\mathbf{Q}_d$ is the covariance matrix of the concentration vector; and $\mathbf{I}$ is the identity matrix. A detailed discussion of the Kalman filtering algorithm for the joint estimation of the model parameters and ICG concentrations in different compartments can be found in [10].

IV. CLINICAL RESULTS

A. Apparatus

In this work, we use the data collected with a continuous wave (CW) NIR imaging apparatus. The apparatus has 16 light sources, which are tungsten bulbs with less than 1 watt of output power. They are located on a circular holder at an equal distance from each other with 22.5 degrees apart. Sixteen detectors, namely, silicon photodiodes, are situated in the same plane. The breast is arranged in a pendular geometry with the source-detector probes gently touching its surface. Figure 2 illustrates the configuration of the apparatus and the configuration of the detectors and the sources in a circular plane. Note that sources and detectors are co-located. The detectors use the same positions as the sources to collect the light originating from one source at a time. Only the signals from the farthest 11 detectors are used in the analysis. For example, when Source 1 is on, the data is collected using Detectors 4 to 14. A band pass filter at 805 nm, the absorption peak of ICG, is placed in front of the sources to select the desired wavelength. A set of data for one source is collected every $\sim$ 500 ms. The total time for a whole scan of the breast including 16 sources and 16 detectors is $\sim$ 8.8 seconds. A more detailed explanation of the apparatus and the data collection procedure can be found in [34].

B. Protocol

Patients with suspicious breast tumors are enrolled for this study. ICG is injected intravenously by bolus with a concentration of 0.25 mg per kg of body weight. Diagnostic information is obtained using biopsy results. Since biopsy modifies the blood volume and blood flow around the tumor region, measurements are made prior to biopsy. Data acquisition started before the injection of ICG and continued for 10 minutes.

C. Tumor Information

Three different patients with different tumor types are included in this study. First case, (Case 1), is fibroadenoma, which corresponds to a mass estimated to be $1-2$ cm in diameter within a breast of 9 cm diameter located at 6-7 o'clock. Second case, (Case 2), is adenocarcinoma corresponding to a tumor estimated to be $2-3$ cm in diameter within a breast of 7.7 cm diameter located at 4-5 o'clock. The third case, (Case 3), is invasive ductal carcinoma, which corresponds to a mass estimated to be 4 by 3 cm located at 6 o'clock. Table 2 describes the tumor information for each patient.

D. ICG Concentration Measurements for Pharmacokinetic Parameter Estimations

Using the CW imager described above, sufficient number of source-detector readings are collected from different angles for each patient. Differential absorption images are reconstructed based on DDOT forward model given in (4). Using the linear relationship (5) between ICG concentration, and absorption coefficient, ICG concentration images are obtained for each case. A sample set of ICG concentration images for the selected time instants are shown in Figure 3, 4, and 5 for Case 1, 2, and 3, respectively. Although only 9 images are displayed, there are around 50 images for each case, each corresponding to a different time instant. Each image is composed of 649 voxels. Note that the ICG concentration images in Figures 3, 4, and 5 represent the bulk ICG concentrations in the tissue, not the ICG concentrations in plasma or in the EES compartments.

We next extract the time course of ICG concentration for each voxel. As an example, Figure 6 shows the time course of ICG concentrations for all three cases for a specific voxel around the tumor region (65th, 276th, 188th voxel for Case 1, Case 2, and Case 3, respectively). We then fit the two-compartment model to each time course data using the EKF framework; and estimate k_{in}, k_{out}, k_{elm}, and the ICG concentrations in plasma and EES. Spatially resolved images of k_{in}, and k_{out} for each case are shown in Figures 8(a)-(b), and 9(a)-(b), 10(a)-(b), respectively. Additionally, we construct ICG concentration images for plasma and the EES compartments. Figures 11-16 show the ICG concentration in plasma and the EES for 3 different time instants for Case 1, 2, and 3, respectively. Our results show that the pharmacokinetic-rates are higher around the tumor region agreeing with the fact that permeability increases around the tumor region due to compromised capillaries of cancer vessels. We also observed that ICG concentrations in plasma and the EES compartments are higher around the tumors agreeing with the hypothesis that around the tumor region ICG may act as a diffusible extravascular flow in leaky capillary of cancer vessels.

Using the brightest spot as a reference point, the *a priori* information on the location, and the size of the tumors, we plot an ellipse (or a circle) to identify the approximate location and size of the tumor in the pharmacokinetic-rate images. Figures 17(a), (b), and (c) present the k_{in} images with approximate tumor location and size for Case 1, 2, and 3, respectively. The consistency of the bright regions in the k_{in} images, and circular/elliptical regions drawn based on the *a priori* information shows that the pharmacokinetic-rate images provide good localization of tumors and size information.

The histograms of k_{in} and k_{out} rates from inside and outside the tumor region (as indicated by circular/elliptical regions) are shown in Figures 18(a), (b), and (c), and Figures 19(a), (b), and (c), respectively. The solid curves in Figures 18 and 19 show the Gaussian fit. Clearly, pharmacokinetic-rate images facilitates analysis of the spatial variation of the pharmacokinetic-rates for breast cancer screening, diagnosis, and staging. The histograms and their Gaussian probability density function fits in Figures 18 and 19 show that the mean and the standard deviation of k_{in} and k_{out} values are different inside and outside the tumor region. Table 3 tabulates the mean values ($\pm$ spatial standard deviation) of the pharmacokinetic-rates from inside and outside the tumor regions for all three cases. The total error (probability of false positive and false negative) using the Bayesian minimum error classifier is given by the overlapping area between the two Gaussian fits [35]. Table 4 tabulates the total Bayesian error. These small error values show that the spatially resolved pharmacokinetics-rate analysis provide good sensitivity and specificity for diagnosis. One other advantage of the pharmacokinetic-rate images is that they allow comparison of rates obtained from healthy and tumor tissue for the patient itself since absolute (mean) values may be quite different from patient to patient. For example, the k_{in} rate for Case 3 obtained outside the tumor region is higher than the k_{in} rate for Case 1 obtained inside the tumor. Thus, using the pharmacokinetic-rates obtained from healthy tissue as a reference to the tumor tissue for the same patient can improve diagnostic accuracy. However, very high mean values of k_{in} and k_{out} may be indicative of the severity of malignancy. The pharmacokinetic rates are higher for Case 3 (invasive ductal carcinoma), for both inside and outside the tumor region as compared to Case 2 (adenocarcinoma). Similarly, the kinetic rates are higher for Case 2 (adenocarcinoma), as compared to Case 1 (fibroadenoma) for both inside and outside the tumor region.

To further understand the value of spatially resolved pharmacokinetic-rate analysis as compared to the bulk pharmacokinetic-rates, we averaged the concentration images spatially, and obtained a bulk concentration value for each time instant. We then extracted a time curve for the bulk ICG concentrations. Figure 7 shows the resulting bulk ICG concentration time curves for each case. Next, we fitted the two-compartment model to the resulting time curves and estimated the bulk pharmacokinetic-rates. Table 5 tabulates the bulk pharmacokinetic-rates for each patient. In Figure 20 and 21, the bulk pharmacokinetic-rates are overlaid on the histograms of the pharmacokinetic-rate images obtained from inside and outside the tumor region. Figures 20 and 21 show the Bayesian minimum error classifier threshold (the value corresponding to the intersection of the histograms) for each case. We see that for Case 1, the bulk rates of k_{in}, and k_{out} are both classified as healthy tissue (outside the tumor region). For Case 2, k_{in} is classified as cancerous tissue (inside the tumor region) and k_{out} is classified as healthy tissue. Similarly for Case 3, k_{in} is classified as healthy and k_{out} is classified as cancerous tissue. Clearly, the bulk rates do not provide consistent diagnosis. On the other hand, spatially resolved pharmacokinetic-rates allow us to compare values from the cancerous region to the healthy background, and provide higher sensitivity, and specificity as shown by the Bayesian total error values in Table IV.

V. CONCLUSION

In this study, we presented a method of forming pharmacokinetic-rate images and reported spatially resolved pharmacokinetic-rates of ICG for three breast cancer patients. To form pharmacokinetic-rate images, we first obtained a sequence of ICG concentration images using the differential diffuse optical tomographic technique. We next employed the two-compartment model, and estimated the pharmacokinetic-rates and concentrations in each compartment for each voxel using the EKF framework. We have shown in our prior work [10] that the EKF framework has a number of advantages in pharmacokinetic-rate estimation, some of which include robust estimation in the presence of measurement noise and dynamic model uncertainties.

The reconstruction of the pharmacokinetic-rate images presented in this work were obtained in separate, decoupled steps. The quantitative accuracy of pharmacokinetic-rate images can be improved by coupling the two reconstruction steps which can better capture the temporal and spatial variations in the ICG concentrations and pharmacokinetic-rates and provide robust reconstruction by eliminating outliers. Such a direct reconstruction technique is currently being developed in our laboratory and will be reported in the near future.

We formed the pharmacokinetic-rate images using the *in vivo* data obtained from three breast cancer patients and studied the value of spatially resolved pharmacokinetic-rates. A clear advantage of pharmacokinetic-rate images is that they allow comparison of the features obtained form the tumor region to the ones obtained from the healthy background for the same patient. This within patient comparison is important because, pharmacokinetic-rates for healthy tissue may be significantly different from patient to patient. Furthermore, spatially resolved images allow us to utilize not only the average rates but also other features related to the geometry of the heterogeneity, and higher order statistics of the pharmacokinetic-rates. The total error for the Bayesian minimum error classifier for all three patients indicates that the pharmacokinetic-rate images can provide high sensitivity and specificity for tumor diagnosis. The comparison of the bulk pharmacokinetic-rates to Bayesian minimum error classifier threshold shows that the diagnostic information provided by the bulk pharmacokinetic-rates is not consistent. Our study shows that the spatially resolved pharmacokinetic-rate analysis is superior to the bulk pharmacokinetic-rates in tumor diagnosis. Our study also indicates that similar results may also hold for tumor staging and treatment monitoring.

REFERENCES

[1] X. Intes, B. Chance, “Non-PET functional imaging techniques: Optical,” *Radio. Clin. of North America*, vol. 43-1, pp. 221-234, 2005.

[2] A. Yodh and B. Chance, “Spectroscopy and imaging with diffusing light,” *Phys. Today*, vol. 48-3, pp. 34-40, 1995.

[3] E. M. Sevick-Muraca, G. Lopez, T. L. Troy, J. S. Reynolds, and C. L. Hutchinson, “Fluorescence and absorption contrast mechanisms for biomedical optical imaging using frequency-domain techniques, *Photochem. Photobiol.*, vol. 66, pp. 55, 1997.

[4] U. Mahmood, C.H. Tung, A. Bogdanov Jr., and R. Weissleder,“Near infrared optical imaging of protease activity for tumor detection,” *Radiology*, vol. 213-3, pp. 866-870, 1999.

[5] S.R. Arridge, “Optical tomography in medical imaging,” *Inverse Probl*. vol. 15, pp. 41-93, 1999.

[6] S. A. Boppart, W. Luo, D. L. Marks, K. W. Singletary, “Optical coherence tomography: feasibility for basic research and image-guided surgery of breast cancer,” *Breast Cancer Res Treat.*, vol. 84-2, pp. 85-97, 2004.

[7] X. Gu, Q. Zhang, L. Larcom, and H. Jiang, “Three-dimensional bioluminescence tomography with model-based reconstruction,” *Optics Express*, vol. 12, Issue 17, pp. 3996-4000, 2004.

[8] D.J. Cuccia, F. Bevilacqua, A. J. Durkin, S. Merritt, B. J. Tromberg, G. Gulsen, H. Yu, J. Wang, and O. Nalcioglu, “*In vivo* quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration,” *Applied Optics*, vol. 42-1, pp. 2940-2950, 2003.

[9] X. Intes, J. Ripoll, Y. Chen, S. Nioka, A. G. Yodh, B. Chance, “*In vivo* continuous-wave optical breast imaging enhanced with Indocyanine Green,” *Med. Phys.* vol. 30-6, pp. 1039-1047, 2003.

[10] B. Alacam, B. Yazici, X. Intes, B. Chance, “Extended Kalman Filtering for the Modeling and Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods”, *Transactions in IEEE Biomedical Engineering*, vol. 53, no. 10, pp. 1861-1871, 2006.

[11] D. Hansen, A. Spence, T. Carski, and M. Berger, “Indocyanine green (ICG) staining and demarcation of tumor margins in a rat glioma model,” *Surg. Neurol.*, vol. 40, pp. 451-456, 1993.

[12] H. Shinohara, A. Tanaka, T. Kitai, N. Yanabu, T. Inomoto, S. Satoh, Hatano, Y. Yamaoka, and K. Hirao, ”Direct measurement of hepatic Indocyanine Green clearance with near-infrared spectroscopy: separate evaluation of uptake and removal,” *Hepatology*, vol. 23, pp. 137-144, 1996.

[13] A. ElDeosky, A. Seifalian, M. Cope, D. Delpy, and B. Davidson, “Experimental study of liver dysfunction evaluated by direct Indocyanine green clearance using near infrared spectroscopy,” *Br. J. Surg.*, vol. 86, pp. 1005-1011, 1999.

[14] P. Vaupel, K. Schlenger, C. Knoop, M. Hockel, “Oxygenation of human tumors: evaluation of tissue oxygen distribution in breast cancers by computerized O_2 tension measurements,” *Cancer Res.*, vol. 51, pp. 33163322, 1991.

[15] C. W. Tornoe, “Grey-Box PK/PD Modeling of Insulin,” M.S. Thesis, June 28, 2002.

[16] D. H. Anderson, *Lecture Notes in Biomathematics: Compartmental Modeling and Tracer Kinetics*, Springer-Verlag, Berlin, 1983.

[17] J. A. Jacquez, *Compartmental Analysis in Biology and Medicine: Kinetics of Distribution of Tracer-labeled Materials*, Elsevier Pub. Co., New York, 1972.

[18] H. Sun, J. M. Collins, T. J. Mangner, O. Muzik, A. Shields, “Imaging the pharmacokinetics of [F-18]FAU in patients with tumors: PET studies,” *Cancer Chemother. Pharmacol*, vol. 57-3, pp. 343-348, 2006.

[19] S. Mussurakis, D. L. Buckley, P. J. Drew, J. N. Fox, P. J. Carleton, L. W. Turnbull, A. Horsman, “Dynamic MR imaging of the breast combined with analysis of contrast agent kinetics in the differentiation of primary breast tumours,” *Clin Radiol.*, vol. 52-7, pp. 516-26, 1997.

[20] M. Y. Su, H. J. Yu, P. M. Carpenter, C. E. McLaren, O. Nalcioglu, “Pharmacokinetic parameters analyzed from mr contrast enhancement kinetics of multiple malignant and benign breast lesions detected in the same patients,” *Technol Cancer Res Treat.*, vol. 4-3, pp. 255-63, 2005.

[21] M. Gurfinkel, A. B. Thompson, W. Ralston, T. L. Troy, A. L. Moore, T. A. Moore, J. D. Gust, D. Tatman, J. S. Reynolds, B. Muggenburg, K. Nikula, R. Pandey, R. H. Mayer, D. J. Hawrysz, and E. M. Sevick-Muraca, “Pharmacokinetics of ICG and HPPH-car for the detection of normal and tumor tissue using fluorescence, near-infrared reflectance imaging: a case study,” *Photochem. Photobiol.*, Vol. 72, pp. 94-102, 2000.

[22] A. B. Milstein, K. J. Webb, and C. A. Bouman, “Estimation of kinetic model parameters in fluorescence optical diffusion tomography,” *J. Opt. Soc. Am.*, vol. 22-7, pp. 1357-1368.

[23] V. Ntziachristos, B. Chance, and A. Yodh,“Differential diffuse optical tomography,” *Opt. Express*, vol. 5, pp. 230-242, 1999.

[24] J. Ripoll, V. Ntziachristos, R. Carminati, and M. Nieto-Vesperinas, “Kirchhoff approximation for diffusive waves,” *Phys. Rev.* vol. 64, pp. 051917, 2001.

[25] J. Ripoll, V. Ntziachristos, J. P. Culver, D. N. Pattanayak, A. G. Yodh, and M. Nieto-Vesperinas, “Recovery of optical parameters in multiplelayered diffusive media: theory and experiments,” *J. Opt. Soc. Am.* vol. A 18, pp. 821830, 2001.

[26] D. Boas, T. Gaudette, and S. Arridge, “Simultaneous imaging and optode calibration with diffuse optical tomography,” *Opt. Express* vol. 8, pp. 263270, 2001.

[27] M. O’Leary,‘Imaging with Diffuse Photon Density Waves,” in PhD Thesis, Dept. Physics & Astronomy, U. of Pennsylvania, May 1996.

[28] P. Hansen and D. OLeary, “The use of the L-curve in the regularization of discrete ill-posed problems,” *SIAM J. Sci. Comput.*, vol. 14, pp. 1487 1503, 1993.

[29] M. L. J. Landsman, G. Kwant, G. A. Mook, W. G. Zijlstra,“Light- absorbing properties, stability, and spectral stabilization of indocyanine green,” *J. Appl. Physiol.*, vol. 40, pp. 575-583, 1976.

[30] C. Chen, *Linear System Theory and Design*, Oxford University Press, New York, 1999.

[31] L. Ljung, “Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems,” *IEEE Tran. Automa. Control*, vol. AC-24, no. 1, pp. 36-50, 1979.

[32] R. Togneri, and L. Deng, “Joint State and Parameter Estimation for a Target-Directed Nonlinear Dynamic System Model,” *IEEE Tran. on Sig. Proc.*, vol. 51, no. 12, pp. 3061-3070, 2003.

[33] L. Nelson, and E. Stear, “The Simultaneous On-Line Estimation of Parameters and States in Linear Systems,” *IEEE Tran. on Auto. Control*, vol.21, pp. 94-98, 1976.

[34] S. Nioka, Y. Yung, M. Schnall, S. Zhao, S. Orel, C. Xie, B. Chance, and S. Solin,“Optical imaging of breast tumor by means of continuous waves,” *Adv. Exp. Med. Biol.*, vol. 411, pp. 227-232, 1997.

[35] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, 1990.

TABLE I

EKF ALGORITHM FOR SIMULTANEOUS ESTIMATION OF STATES AND PARAMETERS.

Initial Conditions	$\begin{bmatrix} \hat{\mathbf{C}}_d(0) \\ \hat{\boldsymbol{\theta}}(0) \end{bmatrix} = \begin{bmatrix} E(\mathbf{C}_d(0)) \\ \hat{\boldsymbol{\theta}}(0) \end{bmatrix}, \quad \mathbf{P}_{0,0} = \begin{bmatrix} Var(\mathbf{C}_d(0)) & 0 \\ 0 & \mathbf{S}_d \end{bmatrix}$
State Estimate Propagation	$\begin{bmatrix} \hat{\mathbf{C}}_d(k\|k-1) \\ \hat{\boldsymbol{\theta}}(k\|k-1) \end{bmatrix} = \begin{bmatrix} \mathbf{K}_d(\hat{\boldsymbol{\theta}}(k-1))\hat{\mathbf{C}}_d(k-1) \\ \hat{\boldsymbol{\theta}}(k-1) \end{bmatrix}$
Error Covariance Propagation	$\mathbf{P}_{k,k-1} = \mathbf{J}_{k-1}\mathbf{P}_{k-1,k-1}\mathbf{J}_{k-1}^T + \begin{bmatrix} \mathbf{Q}_d & 0 \\ 0 & \mathbf{S}_d \end{bmatrix}$
State Estimate Update	$\begin{bmatrix} \hat{\mathbf{C}}_d(k) \\ \hat{\boldsymbol{\theta}}(k) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{C}}_d(k\|k-1) \\ \hat{\boldsymbol{\theta}}(k\|k-1) \end{bmatrix}$ $+\mathbf{G}_k(\mathbf{m}(k) - \mathbf{V}_d(\boldsymbol{\theta}(k\|k-1))\mathbf{C}_d(k\|k-1))$
Error Covariance Update	$\mathbf{P}_{k,k} = [\mathbf{I} - \mathbf{H}_k\boldsymbol{\Lambda}_{k\|k-1}]\mathbf{P}_{k,k-1}$
Kalman Gain	$\mathbf{H}_k = \mathbf{P}_{k,k-1}\boldsymbol{\Lambda}_{k\|k-1}^T[\boldsymbol{\Lambda}_{k\|k-1}\mathbf{P}_{k,k-1}\boldsymbol{\Lambda}_{k\|k-1}^T + \mathbf{R}_d]^{-1}$
Definitions	$\mathbf{J}_k = \begin{bmatrix} \mathbf{K}_d(\hat{\boldsymbol{\theta}}(k)) & \frac{\partial}{\partial\boldsymbol{\theta}}[\mathbf{K}_d(\hat{\boldsymbol{\theta}}(k))\hat{\mathbf{C}}_d(k)] \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \boldsymbol{\Lambda}_{k\|k-1} = \begin{bmatrix} \mathbf{V}_d(\boldsymbol{\theta}(k\|k-1)) \\ 0 \end{bmatrix}^T$

TABLE II

TUMOR INFORMATION FOR EACH PATIENT

	Tumor Type	Tumor Size	Tumor Location
Case 1	Fibroadenoma	1-2 cm	6-7 o'clock
Case 2	Adenocarcinoma	2-3 cm	4-5 o'clock
Case 3	Invasive Ductal Carcinoma	4 by 3 cm	6 o'clock

TABLE III

MEAN AND STANDARD DEVIATION OF PHARMACOKINETIC-RATES FOR INSIDE AND OUTSIDE THE TUMOR REGION

	k_{in} $(sec^{-1}10^{-2})$		k_{out} $(sec^{-1}10^{-2})$		k_{elm} $(sec^{-1}10^{-3})$	
	Inside	Outside	Inside	Outside	Inside	Outside
Case 1	2.14±0.018	0.73±0.011	1.24±0.069	0.43±0.013	4.11±0.057	3.87±0.012
Case 2	2.92±0.076	1.14±0.052	1.58±0.051	0.65±0.036	3.94±0.081	4.12±0.047
Case 3	6.87±0.093	3.06±0.015	4.96±0.048	1.66±0.072	4.49±0.056	4.46±0.081

TABLE IV

TOTAL ERROR (PROBABILITY OF FALSE POSITIVE AND FALSE NEGATIVE) IN TUMOR DIAGNOSIS BASED ON BAYESIAN MINIMUM ERROR CLASSIFIER

	Total Error for k_{in}	Total Error for k_{out}
Case 1	6.67×10^{-2}	13.34×10^{-2}
Case 2	3.23×10^{-2}	13.51×10^{-2}
Case 3	9.09×10^{-2}	3.64×10^{-2}

TABLE V

BULK PHARMACOKINETIC-RATES EXTRACTED FROM THE ENTIRE BREAST TISSUE

	k_{in} $(sec^{-1}10^{-2})$	k_{out} $(sec^{-1}10^{-2})$	k_{elm} $(sec^{-1}10^{-1})$
Case 1	0.84±0.013	0.62±0.017	3.66±0.042
Case 2	2.01±0.022	0.83±0.012	4.01±0.054
Case 3	4.06±0.072	3.36±0.051	4.37±0.052

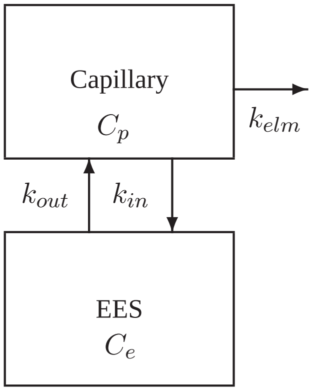

Capillary
C_p
k_{elm}
k_{out}
k_{in}
EES
C_e

Fig. 1.

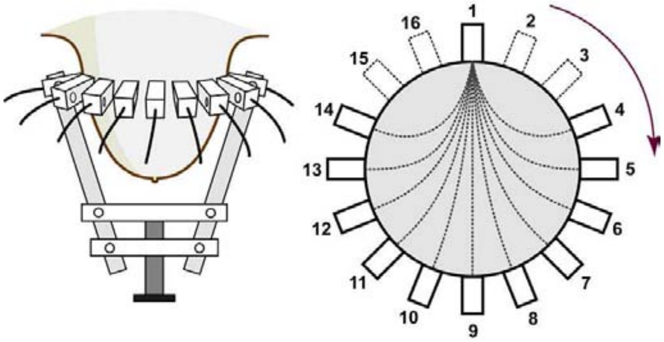


Fig. 2.

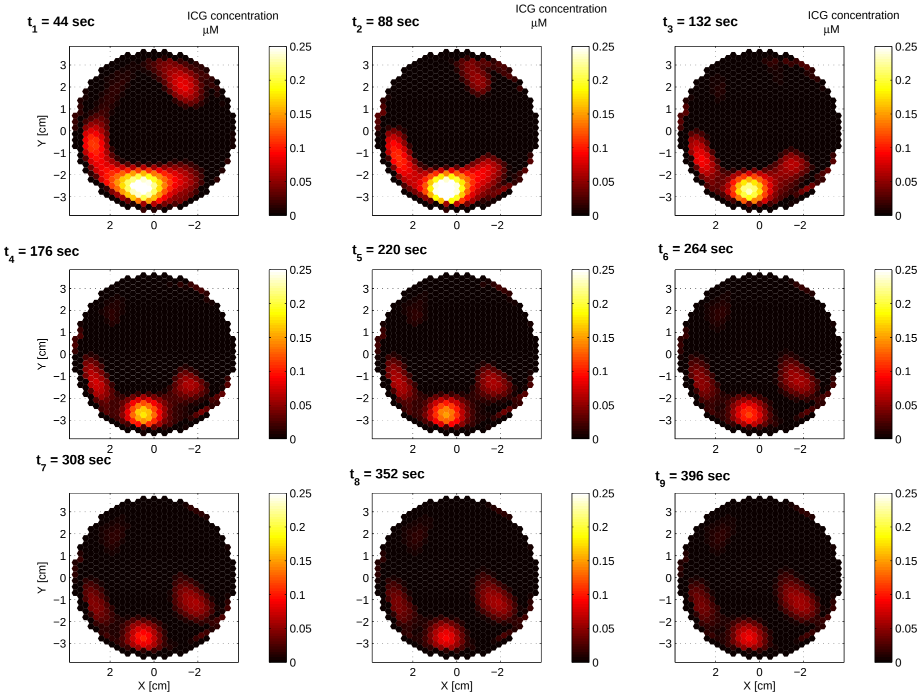
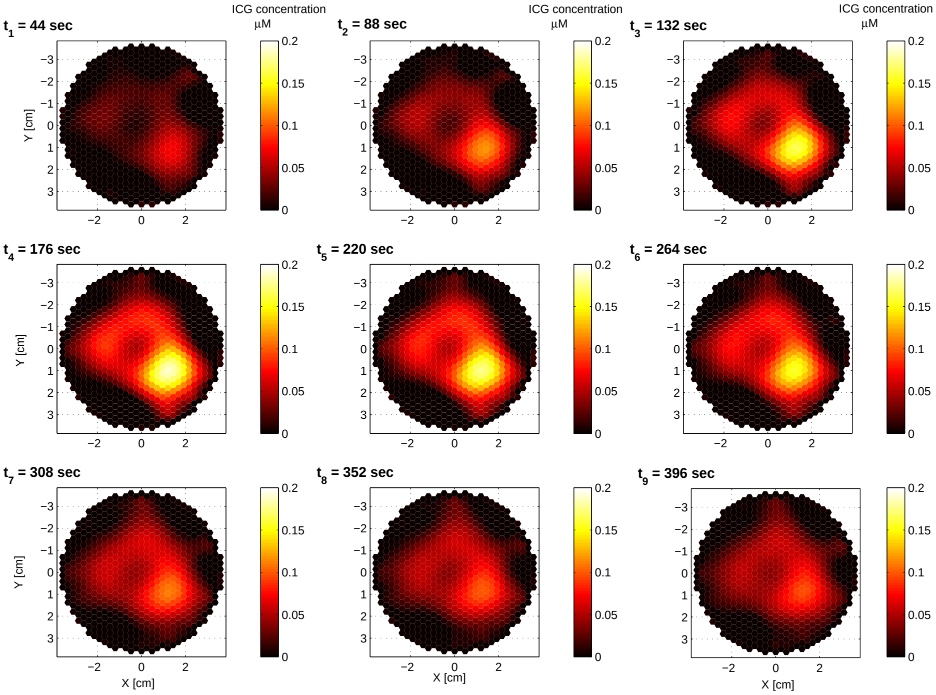

t_1 = 44 sec
t_2 = 88 sec
t_3 = 132 sec
t_4 = 176 sec
t_5 = 220 sec
t_6 = 264 sec
t_7 = 308 sec
t_8 = 352 sec
t_9 = 396 sec
ICG concentration
µM
0.25
0.2
0.15
0.1
0.05
0
Y [cm]
X [cm]
3
2
1
0
−1
−2
−3

Fig. 3.



t_1 = 44 sec
t_2 = 88 sec
t_3 = 132 sec
t_4 = 176 sec
t_5 = 220 sec
t_6 = 264 sec
t_7 = 308 sec
t_8 = 352 sec
t_9 = 396 sec
ICG concentration
μM
0.2
0.15
0.1
0.05
0
Y [cm]
X [cm]
−3
−2
−1
0
1
2
3

Fig. 4.

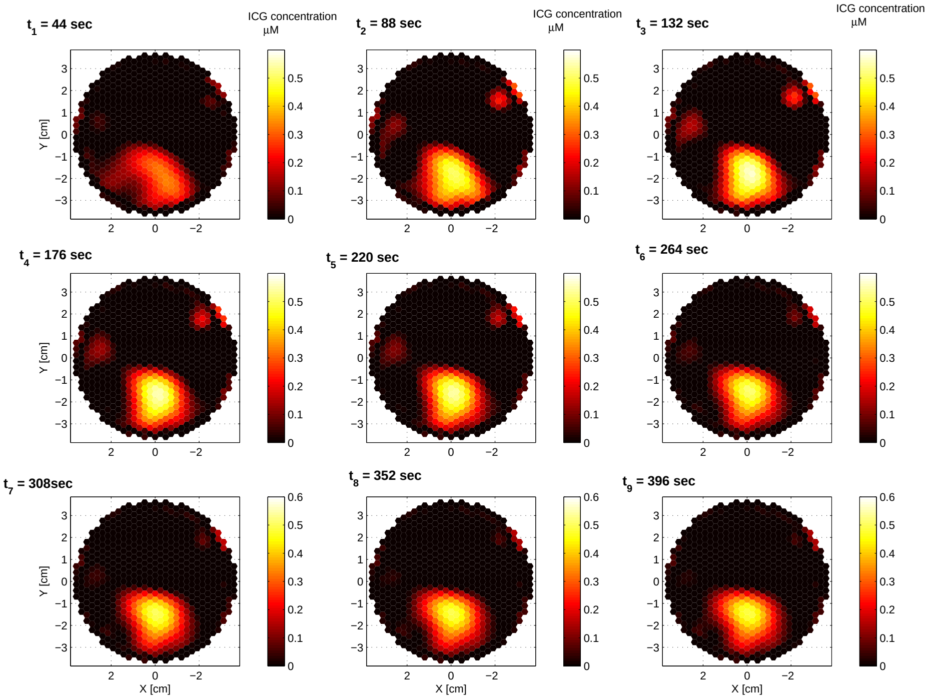
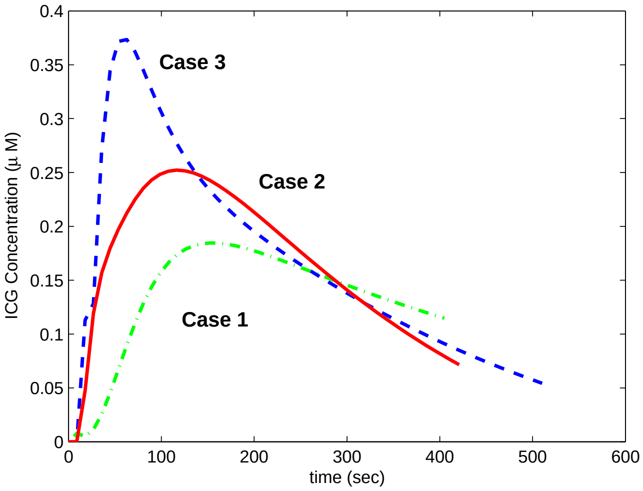
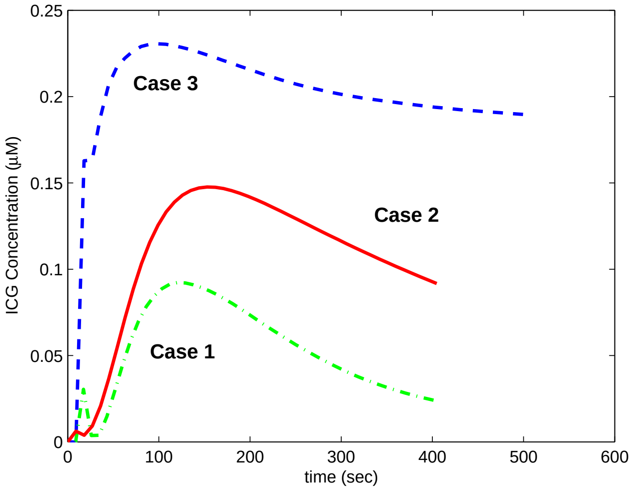

t_1 = 44 sec
t_2 = 88 sec
t_3 = 132 sec
t_4 = 176 sec
t_5 = 220 sec
t_6 = 264 sec
t_7 = 308sec
t_8 = 352 sec
t_9 = 396 sec
ICG concentration
µM
Y [cm]
X [cm]

Fig. 5.



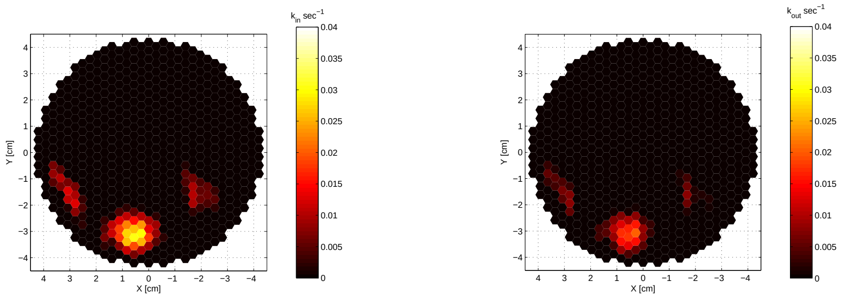
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
ICG Concentration (µ M)
0
100
200
300
400
500
600
time (sec)
Case 3
Case 2
Case 1

Fig. 6.



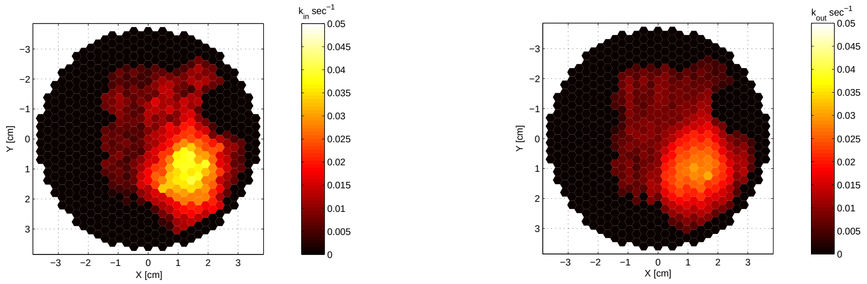
0.25
0.2
0.15
0.1
0.05
0
ICG Concentration (µM)
0
100
200
300
400
500
600
time (sec)
Case 3
Case 2
Case 1

Fig. 7.



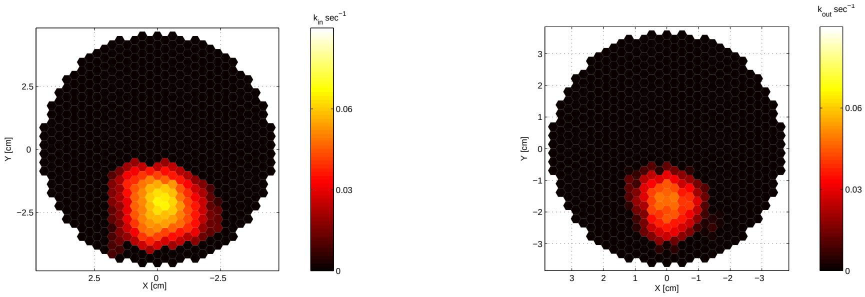
(a) (b)

Fig. 8.



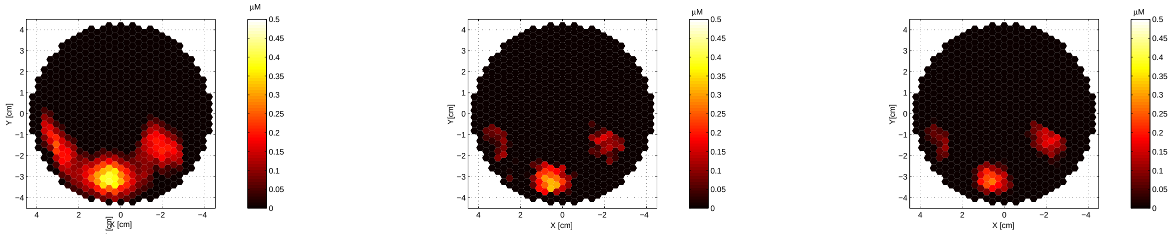
(a) (b)

Fig. 9.



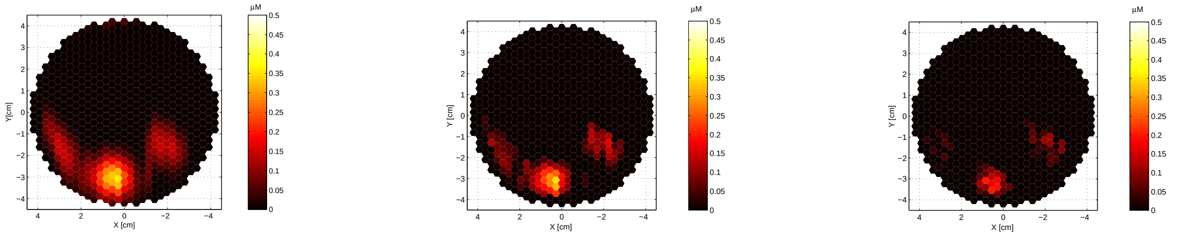
(a) (b)

Fig. 10.



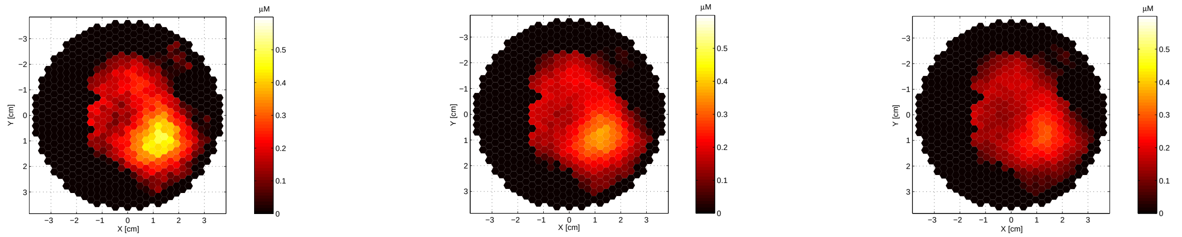
(a) (b) (c)

Fig. 11.



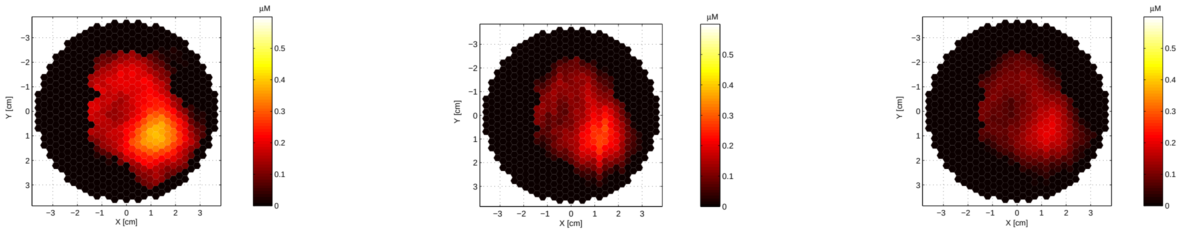
(a) (b) (c)

Fig. 12.



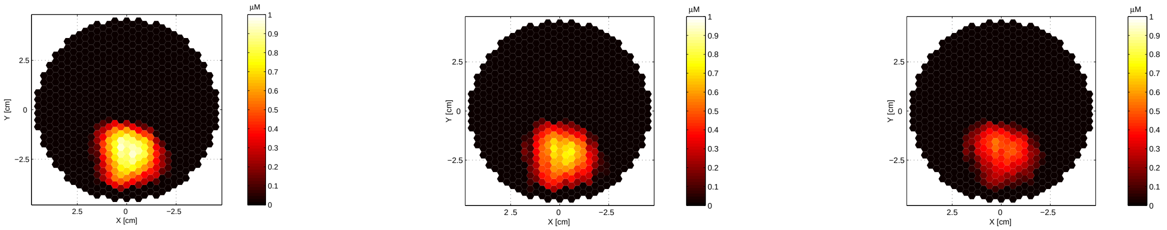
(a) (b) (c)

Fig. 13.



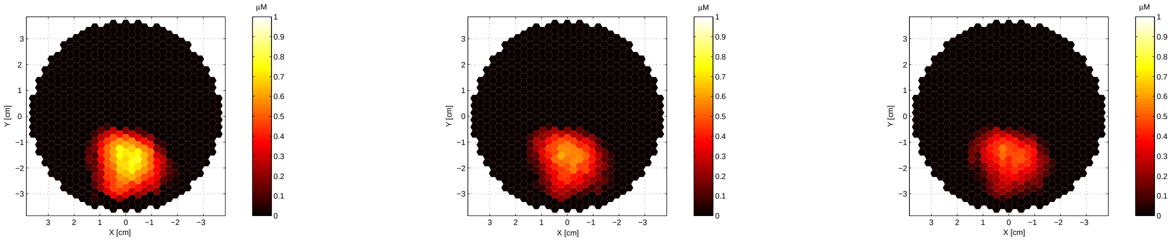
(a) (b) (c)

Fig. 14.



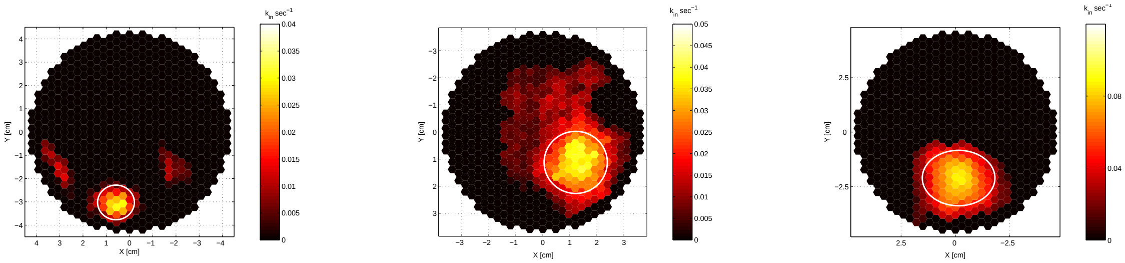
(a) (b) (c)

Fig. 15.



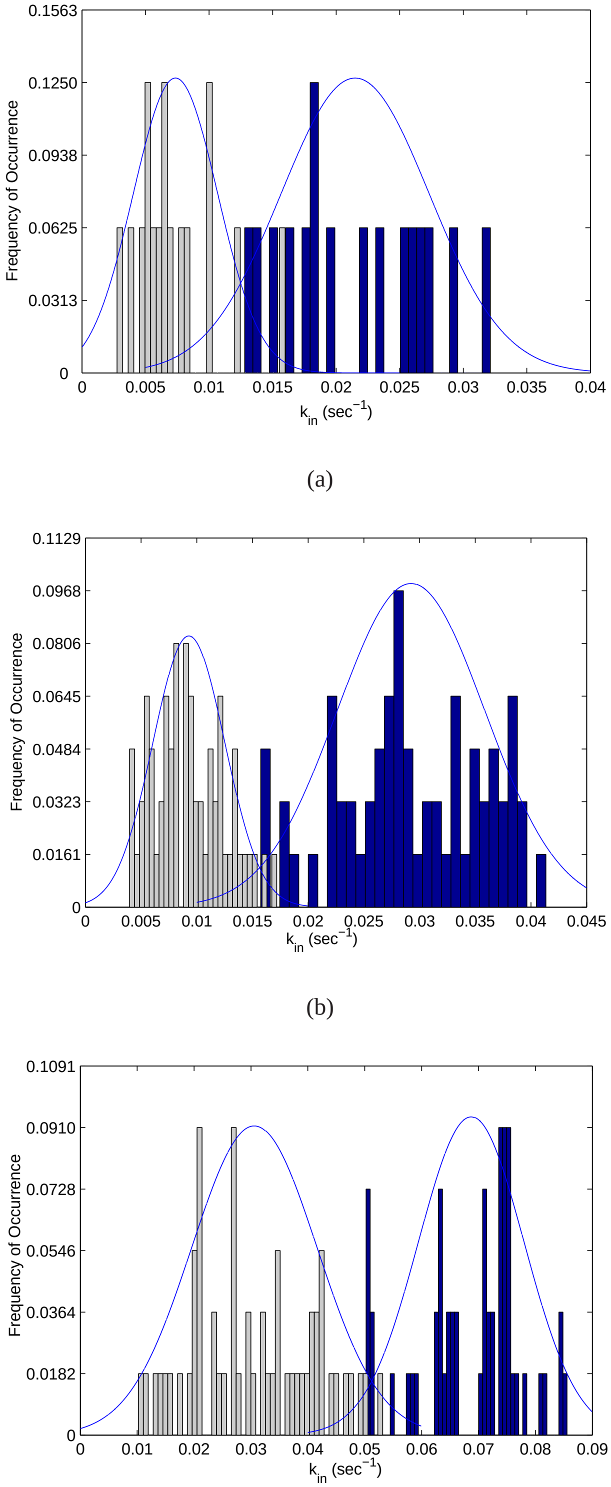
(a) (b) (c)

Fig. 16.



(a) (b) (c)

Fig. 17.


0.1563
0.1250
0.0938
0.0625
0.0313
0
Frequency of Occurrence
0
0.005
0.01
0.015
0.02
0.025
0.03
0.035
0.04
k_{in} (sec^{-1})
0.1129
0.0968
0.0806
0.0645
0.0484
0.0323
0.0161
0
Frequency of Occurrence
0
0.005
0.01
0.015
0.02
0.025
0.03
0.035
0.04
0.045
k_{in} (sec^{-1})
0.1091
0.0910
0.0728
0.0546
0.0364
0.0182
0
Frequency of Occurrence
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
k_{in} (sec^{-1})

(a)

(b)

(c)

Fig. 18.

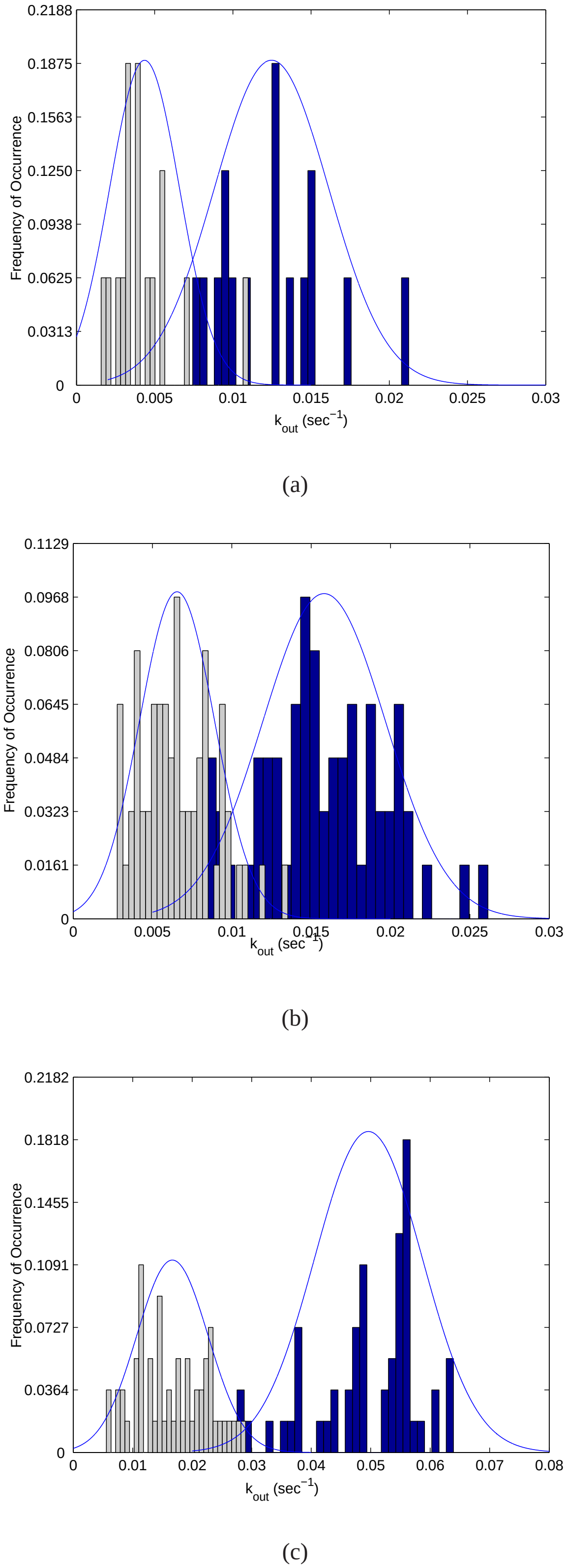
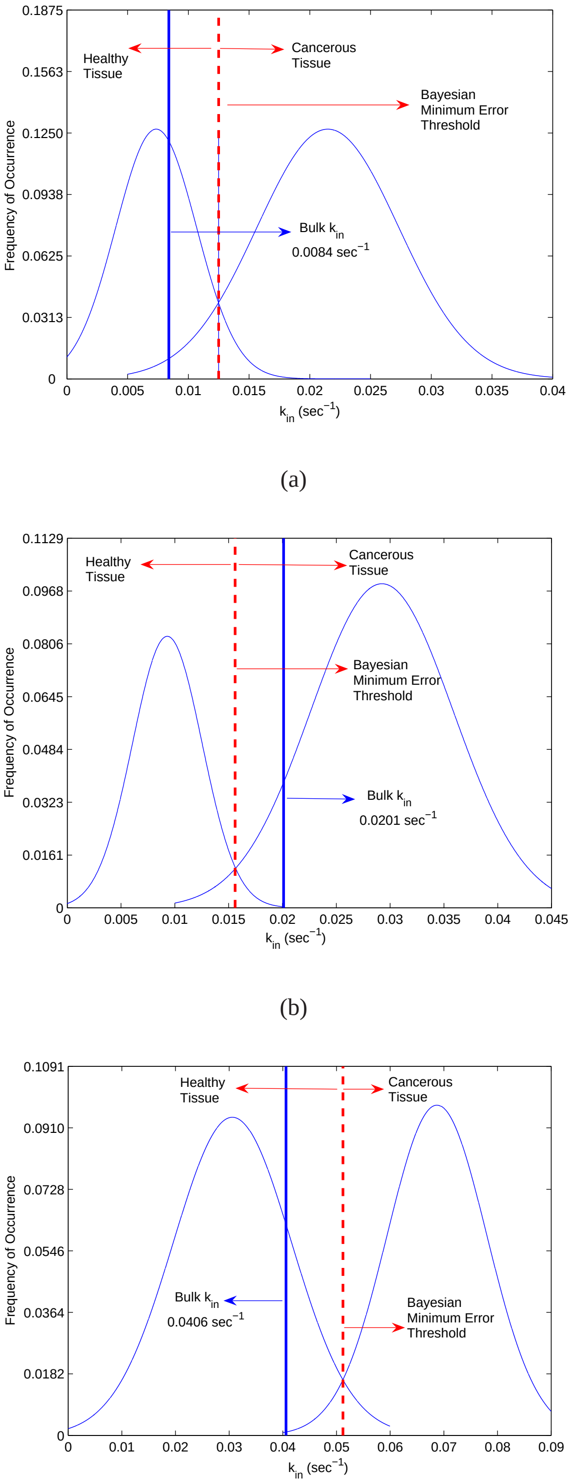

0.2188
0.1875
0.1563
0.1250
0.0938
0.0625
0.0313
0
Frequency of Occurrence
0
0.005
0.01
0.015
0.02
0.025
0.03
k_out (sec^-1)
(a)
0.1129
0.0968
0.0806
0.0645
0.0484
0.0323
0.0161
0
Frequency of Occurrence
0
0.005
0.01
0.015
0.02
0.025
0.03
k_out (sec^-1)
(b)
0.2182
0.1818
0.1455
0.1091
0.0727
0.0364
0
Frequency of Occurrence
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
k_out (sec^-1)
(c)

Fig. 19.

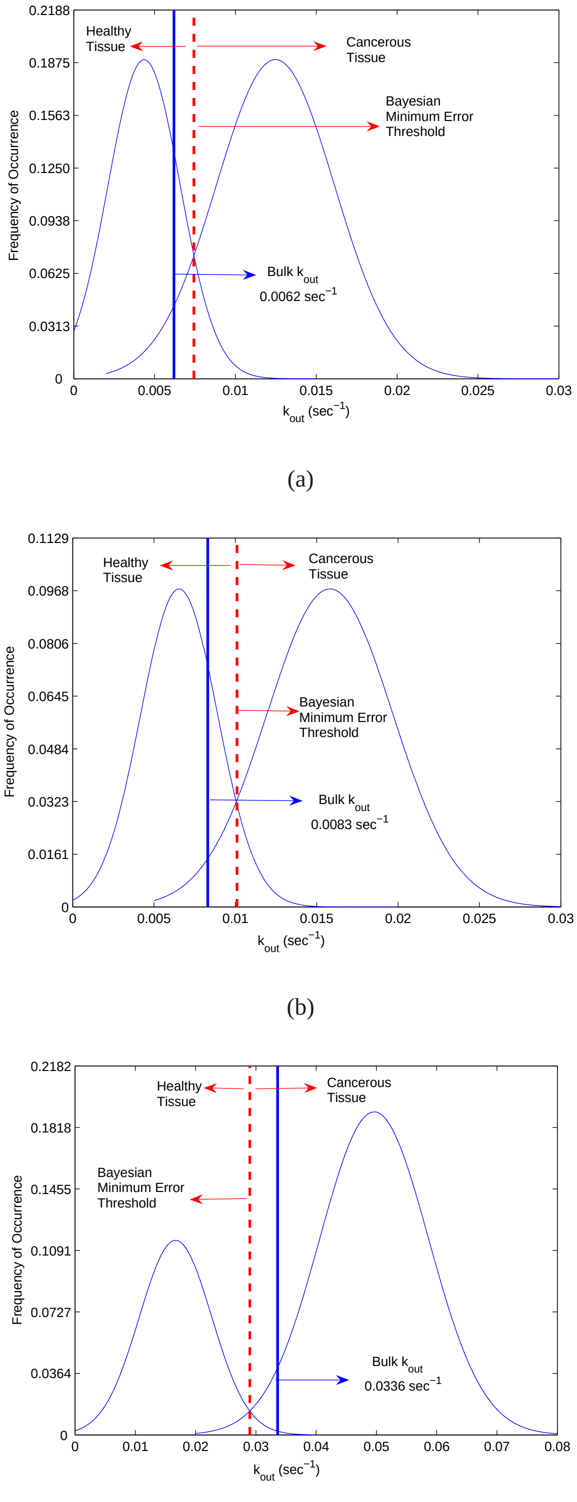

Frequency of Occurrence
0.1875
0.1563
0.1250
0.0938
0.0625
0.0313
0
0
0.005
0.01
0.015
0.02
0.025
0.03
0.035
0.04
k_{in} (sec^{-1})
Healthy Tissue
Cancerous Tissue
Bayesian Minimum Error Threshold
Bulk k_{in}
0.0084 sec^{-1}
Frequency of Occurrence
0.1129
0.0968
0.0806
0.0645
0.0484
0.0323
0.0161
0
0
0.005
0.01
0.015
0.02
0.025
0.03
0.035
0.04
0.045
k_{in} (sec^{-1})
Healthy Tissue
Cancerous Tissue
Bayesian Minimum Error Threshold
Bulk k_{in}
0.0201 sec^{-1}
Frequency of Occurrence
0.1091
0.0910
0.0728
0.0546
0.0364
0.0182
0
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
k_{in} (sec^{-1})
Healthy Tissue
Cancerous Tissue
Bulk k_{in}
0.0406 sec^{-1}
Bayesian Minimum Error Threshold

(a)

(b)

(c)

Fig. 20.


0.2188
0.1875
0.1563
0.1250
0.0938
0.0625
0.0313
0
Frequency of Occurrence
Healthy Tissue
Cancerous Tissue
Bayesian Minimum Error Threshold
Bulk k_{out}
0.0062 sec^{-1}
0
0.005
0.01
0.015
0.02
0.025
0.03
k_{out} (sec^{-1})
0.1129
0.0968
0.0806
0.0645
0.0484
0.0323
0.0161
0
Frequency of Occurrence
Healthy Tissue
Cancerous Tissue
Bayesian Minimum Error Threshold
Bulk k_{out}
0.0083 sec^{-1}
0
0.005
0.01
0.015
0.02
0.025
0.03
k_{out} (sec^{-1})
0.2182
0.1818
0.1455
0.1091
0.0727
0.0364
0
Frequency of Occurrence
Healthy Tissue
Cancerous Tissue
Bayesian Minimum Error Threshold
Bulk k_{out}
0.0336 sec^{-1}
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
k_{out} (sec^{-1})

(a)

(b)

(c)

Fig. 21.

Figure Captions

Fig. 1. Block diagram of the two-compartment model for ICG pharmacokinetics.

Fig. 2. (Left) schematic diagram, (Right) the cut section of the CW NIR imaging apparatus with 16 sources and detectors.

Fig. 3. ICG concentration images for a set of time instants for Case 1.

Fig. 4. ICG concentration images for a set of time instants for Case 2.

Fig. 5. ICG concentration images for a set of time instants for Case 3.

Fig. 6. Time course of ICG concentration curves for a specific voxel, 65th, 276th, 188th voxel for Case 1, Case 2, and Case 3, respectively.

Fig. 7. Averaged time course of ICG concentration curves for Case 1, Case 2, and Case 3, respectively.

Fig. 8. Pharmacokinetic-rate images, (a) k_{in}, and (b) k_{out} for Case 1.

Fig. 9. Pharmacokinetic-rate images, (a) k_{in}, and (b) k_{out} for Case 2.

Fig. 10. Pharmacokinetic-rate images, (a) k_{in}, and (b) k_{out} for Case 3.

Fig. 11. ICG concentration images in plasma for Case 1 for (a) 246.4^{th}, (b) 334.4^{th}, and (c) 422.4^{th} seconds.

Fig. 12. ICG concentration images in the EES for Case 1 for (a) 246.4^{th}, (b) 334.4^{th}, and (c) 422.4^{th} seconds.

Fig. 13. ICG concentration images in plasma for Case 2 for (a) 228.8^{th}, (b) 316.8^{th}, and (c) 404.8^{th} seconds.

Fig. 14. ICG concentration images in the EES for Case 2 for (a) 228.8^{th}, (b) 316.8^{th}, and (c) 404.8^{th} seconds.

Fig. 15. ICG concentration images in the plasma for Case 3 for (a) 246.4^{th}, (b) 378.4^{th}, and (c) 510.4^{th} seconds.

Fig. 16. ICG concentration images in the EES for Case 3 for (a) 246.4^{th}, (b) 378.4^{th}, and (c) 510.4^{th} seconds.

Fig. 17. The k_{in} images with approximate tumor location and size for (a) Case 1, (b) Case 2, and (c) Case 3, respectively. An ellipse/circle identifies the approximate location, and size of the tumor.

Fig. 18. The histograms of k_{in} for (a) Case 1, (b) Case 2, (c), Case 3 from inside (blue) and outside (gray) the tumor region (as indicated by circular/elliptical regions). The solid lines in figures show the Gaussian fit.

Fig. 19. The histograms of k_{out} for (a) Case 1, (b) Case 2, (c), Case 3 from inside (blue) and outside (gray) the tumor region (as indicated by circular/elliptical regions). The solid lines in figures show the Gaussian fit.

Fig. 20. Solid line (blue) shows bulk k_{in} rates for (a) Case 1, (b) Case 2, (c), Case 3 together with the histogram fits. The dashed (red) line indicates the Bayesian minimum error classifier threshold.

Fig. 21. Solid line (blue) shows bulk k_{out} rates for (a) Case 1, (b) Case 2, (c), Case 3 together with the histogram fits. The dashed (red) line indicates the Bayesian minimum error classifier threshold.

Characterization of Breast Tumors with NIR Methods using Optical Indices

Burak Alacam, Birsen Yazici
Dept. of Electrical, Computer, and Systems Engineering,
Rensselaer Polytechnic Institute,
Troy, NY, 12180

Britton Chance, Shoko Nioka
Dept. of Biochemistry and Biophysics,
University of Pennsylvania,
Philadelphia, PA, 19104

Abstract— **This work describes the characterization efficiency of optical properties of breast tumors based on the features obtained using *in vivo* near-infrared (NIR) spectroscopy measurements. Three features, relative blood concentration, oxygen saturation and the size of the tumor, are used to diagnose benign and malignant tumors. The performance of the proposed set of features are evaluated by various classifiers using data acquired from 44 patients with malignant tumors, and 72 patients with benign tumors. The area under the receiver operating characteristics (ROC) curve of the scaled nearest mean classifier (NMSC) using the three features yields a value of 0.91 with a significance level of 0.05. Our results suggest that the features, relative blood concentration, and oxygen saturation can differentiate breast tumors with a relatively high precision.**

I. INTRODUCTION

American cancer society (ACS) estimates that a total of approximately 200,000 new cases of invasive breast cancer occur in women in the United States every year [1]. Currently, there are over 2 million women living in the US who have been diagnosed with and treated for breast cancer. A total of 40,410 women and 470 men are predicted to die from breast cancer in the US during the year 2007 as per ACS estimates [1]. Breast cancer continues to be the leading cancer site among American women. *Early* detection is critical for effective treatment of breast cancer. Patients with tumors 1 cm or less in size have a greater than 90 percent long-term survival [2].

In recent years, there has been considerable interest in near-infrared (NIR) optical spectroscopy and tomography techniques since they provide contrast information that is specific to oxyhemoglobin, deoxyhemoglobin, and water which can potentially be used for *early* detection and diagnosis of breast cancer [3]–[8].

Correct interpretation of the optical indices (i.e. deoxyhemoglobin, oxyhemoglobin, blood volume, water content, scattering, and absorption) obtained by optical spectroscopic/tomographic techniques is also important as well as acquiring them. Several research groups demonstrated that the contrast in optical indices can provide information that allows for better characterization of breast cancer [3]–[5]. In [3], Pogue et al. presented a way to measure and obtain hemoglobin concentration, oxygen saturation, water fraction, scattering power, and scattering amplitude. These indices were then investigated for the differences between healthy and diseased breast tissues. In [4], Grosenick et al. reported on the optical indices, scattering and absorption coefficients, hemoglobin concentration, and blood oxygen saturation obtained using optical measurements. Their results showed that these optical indices can be used to distinguish carcinomas from healthy breast tissues. Recently, Khayat et al. [5] presented characterization results of optical indices, oxyhemoglobin, deoxyhemoglobin, blood volume, lipid and water content, scattering and absorption coefficients, using optical imaging. The results showed the ability of optical imaging to characterize different types of breast lesions.

In this work, we evaluated the characterization efficiency of optical properties of breast tumors using *in vivo* data obtained by near-infrared (NIR) spectroscopy. Our evaluation criteria is based on statistical classification techniques. Three features, namely, relative blood concentration, ΔBV, oxygen saturation, $\Delta Deoxy$, and the size of the tumor, S, were used to characterize benign and malignant tumors. The performance of the proposed set of features were evaluated using various classifiers on 44 patients with malignant tumors, and 72 patients with benign tumors. The area under the receiver operating characteristics (ROC) curve of the scaled nearest mean classifier (NMSC) using the three features yields a value of 0.91 with a significance level of 0.05.

The rest of the paper is organized as follows: In Section II, we present the NIR apparatus, and data protocol, followed by feature extraction and tumor classification. In Section III, we present statistical analysis of clinical data. Section IV summarizes our results.

II. METHODS

A. Apparatus

In this study, a continuous wave (CW) near infrared spectrometer (NIRS) is used [9]. The apparatus includes a probe (Fig. 2). In the center of the probe there is a 3-wavelength light emitting diode (LED). The probe consisted of one multi-wavelength LED as a light source and 8 silicon diodes as detectors. The detectors surround the LED with a 4 cm radius. The light intensity from the detectors was adjusted to be approximately 1 volt and calibrated with a phantom with known absorption and scattering coefficients.

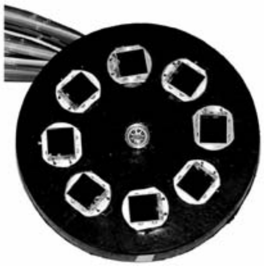


Fig. 1. The NIR probe with a multi-wavelength LED and 8 silicon diodes as detectors.

B. Patients and Protocol

This study includes two centers, namely, the Abramson Family Cancer Research Institute, Department of Radiology of the Hospital of University of Pennsylvania (HUP), and the Department of Gynecology of Leipzig University (DGLU). HUP provided 24 patients with malignant and 64 patients benign tumors. DGLU provided 20 patients with malignant and 6 benign tumors.

The measurements are taken on the breast with tumor. Then, the probe is transferred to the tumor free contralateral breast to include the mirror image location of the suspected cancer. The sensors giving the largest changes with respect to the mirror image position on the contralateral breast are related to the suspected cancer.

C. Feature Extraction

In this study, three features, namely, relative blood concentration, ΔBV, oxygen saturation, $\Delta Deoxy$, and the size of the tumor, S, are used.

The features, ΔBV, and $\Delta Deoxy$ are obtained using

$$\Delta OD = \varepsilon \Delta CL \tag{1}$$

where OD is the optical density, ε is the extinction coefficient, C is blood concentration, L is the mean pathlength of photons, and Δ denotes relative change. Here, $\varepsilon \approx 1$ cm^{-1}, and $L = 4$ cm for a pathlength factor of 5.

Following (1), the relative blood concentration, ΔBV, and the oxygen saturation, $\Delta Deoxy$, can be approximated at two different wavelengths by

$$\Delta BV \propto 0.3\Delta OD_{730} + \Delta OD_{850} \tag{2}$$

$$\Delta Deoxy \propto 1.3\Delta OD_{730} + \Delta OD_{850} \tag{3}$$

where ΔOD_{730}, and ΔOD_{850} denote the relative changes in optical density at 730 nm and 850 nm, respectively.

ΔBV and $\Delta Deoxy$ can also be approximated by

$$\Delta BV \propto \Delta[Hb] + \Delta[HbO_2] \tag{4}$$

$$\Delta Deoxy \propto \Delta[HbO_2] - \Delta[Hb] \tag{5}$$

where $\Delta[Hb]$, and $\Delta[HbO_2]$ denote the relative change in deoxyhemoglobin (Hb) and oxyhemoglobin ($Hb0_2$).

The concentrations of Hb, and HbO_2 in (4) and (5) are calculated by the Beer-Lambert Law given by

$$\Delta OD = log\frac{I_0}{I} \tag{6}$$

where I is light intensity after absorption and scattering, and I_0 is the baseline light intensity obtained from the contralateral breast, using known extinction coefficients of Hb, HbO_2 and differential pathlength factors [10].

Here, it is important to note that, ΔBV, and $\Delta Deoxy$ values are based on a lipid blood oxygen model. Thus the increments of BV and $Deoxy$ are relative to the contralateral breast:

$$\Delta BV = \Delta BV_{tumor} - \Delta BV_{contra} \tag{7}$$

$$\Delta Deoxy = \Delta Deoxy_{tumor} - \Delta Deoxy_{contra} \tag{8}$$

where ΔBV_{tumor}, ΔBV_{contra} are relative blood volume in the tumor breast and the mirror image position of the contralateral breast, respectively, and $\Delta Deoxy_{tumor}$, $\Delta Deoxy_{contra}$ are relative oxygen saturation in the tumor breast and the mirror image position of the contralateral breast, respectively.

D. Feature Analysis and Tumor Classification

In this subsection, we present the set of tumor classification features, and the malignancy differentiation criteria. F1 denotes ΔBV, F2 denotes $\Delta Deoxy$, and F3 denotes, S, size of the tumor. We evaluate the malignancy differentiation capability of the individual features and various combinations of the these features using a set of classifiers, namely, k-nearest neighbor classifier (KNNC), Parzen density based classifier (PAR), automatic neural network classifier (NEURC), normal densities based linear classifier (LDC), nearest mean classifier (NMC), scaled nearest mean classifier (NMSC), normal densities based quadratic classifier (QDC), uncorrelated normal densities based quadratic classifier (UDC). The more details information on these classifiers can be found in [12].

We evaluated the malignancy differentiation capability of the following individual and combined features:

F1: ΔBV
F2 : $\Delta Deoxy$
F3 : Tumor Size (S)
F1-F2: ΔBV and $\Delta Deoxy$
F1-F2-F3: ΔBV, $\Delta Deoxy$, and S

III. STATISTICAL ANALYSIS OF CLINICAL DATA

The evaluation is based on receiver operating characteristics (ROC) methodology. The ROC curve is obtained by plotting the probability of false positive rate versus the probability of detection. The evaluation of classification method is done using area under the ROC curve (AUC). First, we evaluated the classification performance of all three features. Table I presents the AUC values for 8 different classifiers for all three features. The NMSC has the best performance in terms of classification with a AUC value of 0.9098 followed by the Parzen classifier with a AUC value of 0.9041.

TABLE I

AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F1-F2-F3: ΔBV, $\Delta Deoxy$, S

Type	**NMSC**	PAR	LDC	UDC	NEURC	QDC	NMC	KNNC
AUC	**0.9098**	0.9041	0.9017	0.8984	0.8864	0.8843	0.8807	0.8752

TABLE II

AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F1-F2: ΔBV, $\Delta Deoxy$

Type	**NMSC**	PAR	LDC	UDC	NEURC	QDC	NMC	KNNC
AUC	**0.9001**	0.8993	0.8930	0.8908	0.8992	0.8821	0.8782	0.8645

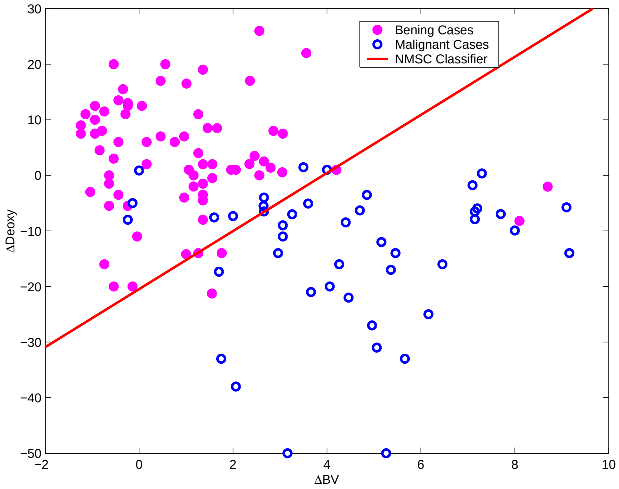


Fig. 2. Scaled Nearest Mean Classifier and F1-F2 2-D data clustering.

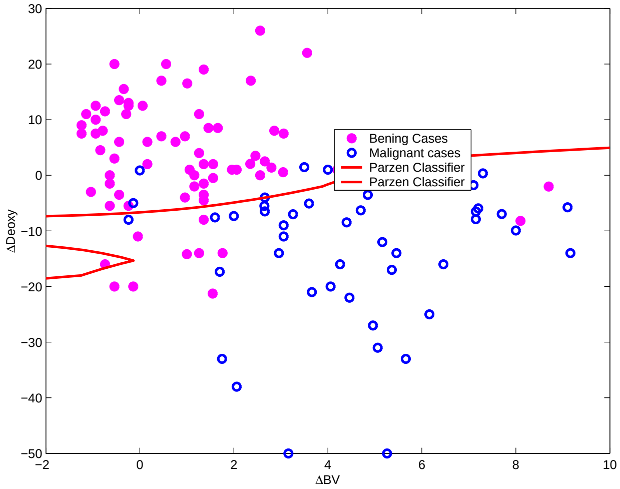


Fig. 3. Parzen and F1-F2 2-D data clustering.

Next, we evaluated the performance of the two features measured by NIR spectroscopy. Table II presents the AUC values for 8 different classifiers for features ΔBV and $\Delta Deoxy$. Again, the NMSC performed the best in terms of classification with a AUC value of 0.9001. Finally, we evaluated the individual classification performances of the three features. Table III presents the AUC values for 8 different classifiers for the feature ΔBV. The NMC has the best performance in terms of classification with a AUC value of 0.8832. Table IV presents the results 8 different classifiers for the feature $\Delta Deoxy$. The NMC has the best performance in terms of classification with a AUC value of 0.879. Table V presents the results 8 different classifiers for the feature S. The QDC has the best performance in terms of classification with a AUC value of 0.5612.

As it can be seen from Tables I, and II, the best performing feature set is the combination of the three features. We can also conclude from Table V that, the tumor size can not be used to differentiate healthy and diseased tissues with an AUC value of around 0.5. However, the combination set of optical indices, obtained using optical measurements, can differentiate breast tumors with a relatively high precision with a AUC value of 0.9. Similarly, optical indices, ΔBV and $\Delta Deoxy$, also performed well with AUC values of 0.883 and 0.879, respectively.

Figures 2, and 3 show the distribution of features ΔBV, and $\Delta Deoxy$ extracted from benign and malignant tumors. The thresholds were computed using scaled nearest mean classifier and Parzen classifiers. Figure 4 presents, the ROC curves for all three features, and the best two features, namely, $\Delta Deoxy$ and ΔBV. The observed area under the ROC curve for F1-F2-F3, and F1-F2 are 0.9098, and 0.9001, respectively. Figure 5 presents the ROC curves for individual features F1, and F2 using the nearest mean classifier. The observed area under the ROC curve for F1, and F2 are 0.8832 and 0.8790, respectively.

IV. CONCLUSION

In this work, we evaluated the characterization efficiency NIR optical spectroscopy using three features, relative blood concentration, oxygen saturation, and the size of the tumor. The characterization of malignant and benign tumors are evaluated using different classifiers. Our results suggest that the relative blood concentration, and oxygen saturation has potential to differentiate malignant and benign breast tumors with a relatively high accuracy. This set of features can potentially be incorporated into a diagnostic systems to aid physicians for breast cancer diagnosis. In the near future, we will incorporate additional features to the current feature set. We plan to analyze the new set of features using different

TABLE III
AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F1: ΔBV

Type	NMSC	PAR	LDC	UDC	NEURC	QDC	**NMC**	KNNC
AUC	0.8817	0.8764	0.8807	0.8779	0.8513	0.8778	**0.8832**	0.8302

TABLE IV
AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F2: $\Delta Deoxy$

Type	NMSC	PAR	LDC	UDC	NEURC	QDC	**NMC**	KNNC
AUC	0.8787	0.8764	0.8776	0.8711	0.8491	0.8613	**0.8790**	0.8331

TABLE V
AUC VALUES FOR DIFFERENT CLASSIFIERS FOR F3: S

Type	NMSC	PAR	LDC	UDC	NEURC	**QDC**	NMC	KNNC
AUC	0.5123	0.5292	0.4782	0.5429	0.5382	**0.5612**	0.5112	0.4827

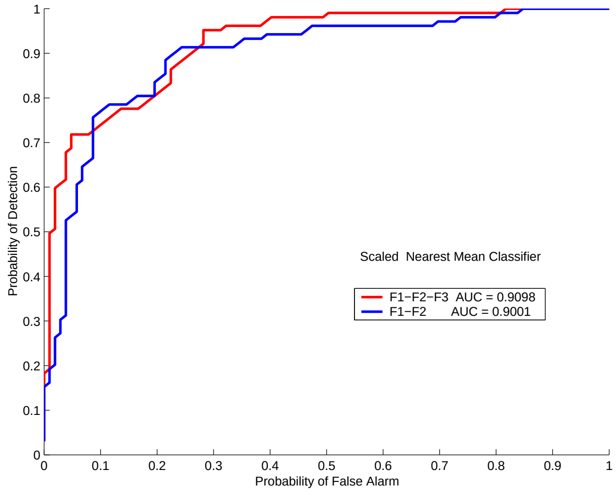


Fig. 4. ROC curves for F1-F2-F3 and F1-F2 using NMSC Classifier.

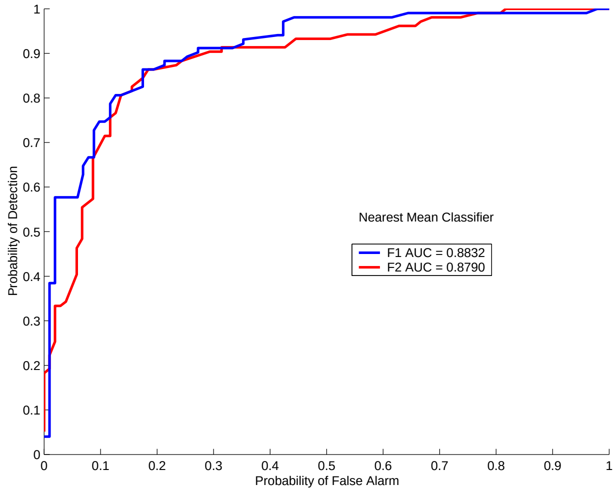


Fig. 5. ROC curves for F1and F2 using NMC Classifier.

sampling techniques, i.e. hold-out, leave-one-out, resubstitution techniques, and further investigate the characterization efficiency optical features. We also plan to compare the ROC performance of the optical features, F1 and F2, with that of F3, using a hypothesis testing method based on the AUC statistics.

REFERENCES

[1] American cancer society, statistics for 2006: http://www.cancer.org/

[2] Report of the Joint Working Group on Quantitative In Vivo Functional Imaging in Oncology, Sponsored by the *U.S. Public Health Services Office on Womens Health* and *National Cancer Institute*, 1999.

[3] B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, S. Srinivasan, X. Song, T. D. Tosteson, S. P. Poplack, K. D. Paulsen, “Characterization of hemoglobin, water, and NIR scattering in breast tissue: analysis of intersubject variability and menstrual cycle changes,” *Journal of Biomedical Optics*, Vol. 9(3), pp. 541552, 2004.

[4] D. Grosenick, H. Wabnitz, K T. Moesta, J. Mucke, M. Mller, C. Stroszczynski, J. Stel, B. Wassermann, P. M. Schlag, and H. Rinneberg, “Concentration and oxygen saturation of haemoglobin of 50 breast tumours determined by time-domain optical mammography,” *Phys. Med. Biol.* Vol. 49 No. 7, pp. 1165-1181, 2004.

[5] M. Khayat, Z. Ichalalene, N. Mincu, F. Leblond, O. Guilman, and S. Djeziri, “Optical tomography as adjunct to x-ray mammography: methods and results,” *Proc. of SPIE*, Vol. 6431 64310F-1, 2007.

[6] B. J. Tromberg, N. Shah, R. Lanning, A. Cerussi, J. Espinoza, T. Pham, L. Svaasand, and J. Butler, Non-invasive in vivo characterization of breast tumors using photon migration spectroscopy, Neoplasia Vol. 2 (12), pp. 2640, 2000.

[7] V. Ntziachristos, B. Chance, “Probing physiology and molecular function using optical imaging: applications to breast cancer,” *Breast Cancer Res.*, Vol. 3, pp. 41-46, 2001.

[8] P. Vaupel, A. Mayer, S. Briest, and M. Hockel, “Oxygenation gain factor: a novel parameter characterizing the association between hemoglobin level and the oxygenation status of breast cancers,” *Cancer Res.*, Vol. 63, pp. 76347637, 2003.

[9] B. Chance, S. Nioka, J. Zhang, E. F. Conant, E. Hwang, S. Briest, S. G. Orel, M. D. Schnall, B. J. Czerniecki, “Breast cancer detection based on incremental biochemical and physiological properties of breast cancers: a six-year, two-site study,” *Academic Radiology*, Vol. 12, Issue 8, pp. 925-933, 2005.

[10] S. Fantini, D. Hueber, M. A. Franceschini, E. Gratton, W. Rosenfeld, P. G. Stubblefield, D. Maulik, M. R. Stankovic, “Non-invasive optical monitoring of the newborn piglet brain using continuous-wave and frequency-domain spectroscopy,” *Phys. Med. Biol.*, Vol. 44, pp. 1543-1563, 1999.

[11] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, New York:Academic Press, 1990.

[12] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, New York: Wiley-Interscience, 2000.

Reconstruction of Spatially Resolved Pharmacokinetic Rate Images of Fluorescence Agents in FDOT

Burak Alacam, Birsen Yazici
Ahmed Serdaroglu
Dept. of Electrical, Computer, and
Systems Engineering
Rensselaer Polytechnic Institute
Troy, New York, 19102

Xavier Intes
ART Inc.
Saint-Laurent, Quebec
H4S 2A4, Canada

Britton Chance, Shoko Nioka
Department of Biochemistry and Biophysics,
University of Pennsylvania,
Philadelphia, PA, 19104

Abstract— **In this work, we propose a method to reconstruct spatially resolved pharmacokinetic rate images of fluorescence agents directly from the boundary photon flux measurements. We use a compartmental modeling scheme to model the pharmacokinetics of fluorescence agents. We coupled this model with the fluorescence diffuse optical tomography (FDOT) forward model to form a state space model which is then iteratively solved by extended Kalman filtering (EKF) algorithm. As an example, we used a two-compartment model for indocyanine green (ICG) pharmacokinetics. To validate the proposed method, we tested our approach using a simulation study. Reconstructed pharmacokinetic rate images with correct localization of heterogeneities and high reconstruction accuracy show that the proposed algorithm can be used for tumor detection, tumor diagnosis, drug delivery and feasibility studies.**

I. INTRODUCTION

Development of fluorescence agents that emit light upon excitation lead to a new imaging technique, called fluorescence diffuse optical tomography (FDOT), whose theory is a straightforward extension of diffuse optical tomography. In FDOT, near infrared (NIR) excitable exogenous fluorescence agents, which are externally injected into the tissue are investigated in terms of their optical parameters [1]. These agents enhance the sensitivity of detection hence increase the optical contrast. FDOT has the ability to characterize functional parameters of tissue such as scattering and absorption, which can be used to derive pharmacokinetics of fluorescence agents.

The analysis of pharmacokinetics of fluorescence agents is a potential means for tumor detection, diagnosis, drug delivery and feasibility studies [2]. One approach to pharmacokinetic analysis is the compartmental modeling [6]. In this method, a region of interest consists of a number of compartments, generally representing a volume of similar tissues into which the fluorescence agent is distributed. The concentration changes in a specific compartment are modeled as a result of the exchange of fluorescence agents between connected compartments. These changes are modeled by a collection of coupled ordinary differential equations (ODE); each equation describing the time change dictated by the biological laws that govern the concentration exchanges between the interacting compartments. Coefficients of the ODE's are the pharmacokinetic rates that represent rates of exchange between different compartments.

A number of research groups reported compartmental modeling of fluorescence agents for tumor diagnosis in animal and human subjects [3]–[5]. In these studies, bulk values that represent pharmacokinetic rates either over the entire imaging domain or a large volume is used for tumor detection/diagnosis. In such cases, spatially resolved representation of pharmacokinetic rates may increase specificity and sensitivity compared to average bulk rates. Moreover, there are studies in literature showing that the spatially varying pharmacokinetics rates are quantitatively different inside and outside the tumor region [7]–[9].

In this paper, we proposed a compartmental modeling scheme for fluorescence agent pharmacokinetics, and reconstructed spatially resolved pharmacokinetic rate images directly from the boundary photon flux measurements. We also reconstructed the spatially resolved fluorescence agent concentration images inside different compartments. In our algorithm, we used a finite element method (FEM) [10] as a forward solver and an extended Kalman filter (EKF) algorithm [3] as an inverse solver. We tested our approach using a simulated data set based on a two-compartment model for ICG pharmacokinetics. For the simulation studies, we generated a set of time series data from a domain with tissue-like characteristics using FEM algorithm to solve the coupled diffusion equations with Robin type boundary conditions. Reconstructed pharmacokinetic images with correct localization of heterogeneities and high reconstruction accuracy show that the proposed algorithm can be useful for tumor diagnosis, drug delivery and feasibility studies.

The rest of the paper is organized as follows: In Section II, we present pharmacokinetic analysis and compartmental modeling for fluorescence agents. In Section III, we present FDOT forward problem. Section IV describes the inverse problem and the EKF algorithm for direct reconstruction of pharmacokinetic parameters. In Section V, we present the simulation results. Section VI summarizes our results and conclusion.

II. PHARMACOKINETIC MODELING AND COMPARTMENTAL MODEL ANALYSIS

In general, the continuous time state-space representation for an n-compartment model is given by

$$\dot{\mathbf{C}}(t) = \mathbf{K}(\boldsymbol{\alpha}_n)\mathbf{C}(t) \tag{1}$$

$$m(t) = \mathbf{V}(\boldsymbol{\alpha}_n)\mathbf{C}(t) \tag{2}$$

where $\mathbf{C}(t)$ denotes the concentration vector whose elements are the concentrations of the fluorescence agents in different compartments, and $\dot{\mathbf{C}}(t)$ denotes its time derivative. $m(t)$ is the bulk fluorescence concentration, n is the number of compartments, $\boldsymbol{\alpha}_n$ is the parameter vector whose elements are the pharmacokinetic rate constants and volume fractions. $\mathbf{K}(\boldsymbol{\alpha}_n)$ is the system matrix with entries being the pharmacokinetic rates, and $\mathbf{V}(\boldsymbol{\alpha}_n)$ is the vector containing the volume fractions [6].

Although the formulation for the direct reconstruction of pharmacokinetic rate images can be applicable for n-compartment models, here the two-compartment model for ICG pharmacokinetics will be our running example [3], [4]. Based on a two-compartment model for ICG kinetics, as shown in Figure 1, the ICG transition between two compartments can be modeled by two coupled differential equations:

$$\frac{dC_e(t)}{dt} = -k_{out}C_e(t) + k_{in}C_p(t), \tag{3}$$

$$\frac{dC_p(t)}{dt} = -(k_{in} + k_{elm})C_p(t) + k_{out}C_e(t). \tag{4}$$

Here, $\mathbf{C}(t)$ in (1) and (2) includes, $C_p(t)$ and $C_e(t)$, representing the ICG concentrations in the plasma and extracellular extravascular space (EES), respectively. The parameter vector, $\boldsymbol{\alpha}_2$, includes, k_{in}, k_{out}, k_{elm}, v_p, and v_e,;

$$\boldsymbol{\alpha}_2 = [k_{in} \quad k_{out} \quad k_{elm} \quad v_p \quad v_e]^T, \tag{5}$$

where k_{in} and k_{out} are the pharmacokinetic rates that govern the leakage into and the drainage out of the EES, k_{elm} describes the ICG elimination from the body through kidneys and livers, and v_p and v_e are the plasma and EES volume fractions, respectively.

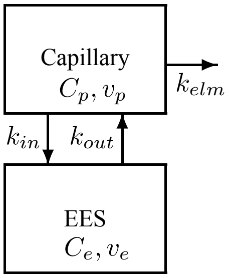


Fig. 1. Block diagram of the two-compartment model for ICG pharmacokinetics.

Here, to obtain 2-D images of pharmacokinetic rates and fluorescence agent concentrations in different compartments, we extend the compartmental model equations (1) and (2) to spatially resolved model as follows:

$$\begin{bmatrix} \dot{\mathbf{C}}(r_1,t) \\ \vdots \\ \dot{\mathbf{C}}(r_N,t) \end{bmatrix} = \begin{bmatrix} \mathbf{K}(\boldsymbol{\alpha}_n(r_1)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{K}(\boldsymbol{\alpha}_n(r_N)) \end{bmatrix} \begin{bmatrix} \mathbf{C}(r_1,t) \\ \vdots \\ \mathbf{C}(r_N,t) \end{bmatrix} + \boldsymbol{\omega}(\mathbf{r},t) \tag{6}$$

where $\boldsymbol{\omega}(\mathbf{r},t)$ is uncorrelated zero mean Gaussian processes with covariance matrix $\mathbf{Q}$, representing the small deviations resulting from model mismatch. $\mathbf{r} = [r_1 \quad r_2....r_N]$, and r_j is the location of the jth voxel for $j = 1, 2, 3...N$, N being the number of total voxels in the discritized domain.

$$\begin{bmatrix} m(r_1,t) \\ \vdots \\ m(r_N,t) \end{bmatrix} = \begin{bmatrix} \mathbf{V}(\boldsymbol{\alpha}_n(r_1)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{V}(\boldsymbol{\alpha}_n(r_N)) \end{bmatrix} \begin{bmatrix} \mathbf{C}(r_1,t) \\ \vdots \\ \mathbf{C}(r_N,t) \end{bmatrix} + \boldsymbol{\eta}(\mathbf{r},t) \tag{7}$$

where $\boldsymbol{\eta}(\mathbf{r},t)$ is a zero mean Gaussian process with covariance matrix $\mathbf{R}$ presenting the noise in the measurements.

The implicit form of (6) and (7) are given by:

$$\dot{\mathbf{C}}(\mathbf{r},t) = \mathbf{K}(\boldsymbol{\alpha}_n(\mathbf{r}))\mathbf{C}(\mathbf{r},t) + \boldsymbol{\omega}(\mathbf{r},t) \tag{8}$$

$$\mathbf{m}(\mathbf{r},t) = \mathbf{V}(\boldsymbol{\alpha}_n(\mathbf{r}))\mathbf{C}(\mathbf{r},t) + \boldsymbol{\eta}(\mathbf{r},t) \tag{9}$$

The quantity we wish to reconstruct is the spatially varying pharmacokinetic rate parameters instead of the absorption coefficient of the florescence agents as commonly the case in FDOT [1]. To do this, we first need to develop a mapping which relates the boundary flux measurements to the bulk florescence agent concentrations in tissue. This mapping can be obtained by using FDOT forward model which will be explained in the next section.

III. FLUORESCENCE DIFFUSE OPTICAL TOMOGRAPHY

We used a coupled system of diffusion equations to model fluorescence light propagation in tissue [1]. Based on the coupled diffusion equations, the forward model for FDOT can be expressed as:

$$\boldsymbol{\Psi}(\mathbf{r},t) = f(\boldsymbol{\mu}_{af}(\mathbf{r},t)), \tag{10}$$

where $\boldsymbol{\mu}_{af}(\mathbf{r},t) = [\mu_{af}(r_1,t), \cdots, \mu_{af}(r_N,t)]$ is the time dependent absorption coefficient of the fluorescence agent, $\boldsymbol{\Psi}(\mathbf{r},t) = [\Psi(r_1,t), \cdots, \Psi(r_N,t)]$ is the time-dependent boundary flux measurements, and f is a nonlinear function defined by the coupled diffusion equation. Under the assumption that the exogenous optical properties has no effect on endogenous optical properties, (10) can be linearized to obtain

$$\boldsymbol{\Psi}(\mathbf{r},t) = \mathbf{W}\boldsymbol{\mu}_{af}(\mathbf{r},t), \tag{11}$$

where $\mathbf{W}$ is the weight matrix which maps the absorption coefficients to the boundary measurements obtained using the FEM algorithm [10].

The absorption coefficient of the fluorescence agents is related to the bulk fluorescence agent concentration as follows:

$$\boldsymbol{\mu}_{af}(\mathbf{r},t) = \epsilon\mathbf{m}(\mathbf{r},t) \tag{12}$$

where ϵ is the extinction coefficient of the fluorescence agent at the excitation wavelength.

A. Pharmacokinetic Rate to Measurement Map for Fluorescence Agents

To derive a complete formulation to reconstruct the pharmacokinetic parameters and concentrations in different compartments using the photon flux measurements, we combined the FDOT forward problem equations, (10) or (11) with the compartmental model equations.

Combining (9) and (10), with the knowledge of the linear relationship between the absorption coefficient and the bulk fluorescence concentration, (12), nonlinear FDOT forward problem is related to the bulk fluorescence agent concentration as:

$$\mathbf{\Psi}(\mathbf{r},t) = g(\mathbf{m}(\mathbf{r},t)) \sim g(\mathbf{V}(\boldsymbol{\alpha}_n(\mathbf{r}))\mathbf{C}(\mathbf{r},t)) \tag{13}$$

where g is the non-linear function which relates the bulk fluorescence agent concentrations to the boundary flux measurements.

The linearized version of (11) is given as:

$$\begin{aligned} \mathbf{\Psi}(\mathbf{r},t) &= \epsilon\mathbf{W}\mathbf{m}(\mathbf{r},t) \\ &= \epsilon\mathbf{W}\mathbf{V}(\boldsymbol{\alpha}_n(\mathbf{r}))\mathbf{C}(\mathbf{r},t) + \epsilon\mathbf{W}\boldsymbol{\eta}(\mathbf{r},t) \\ &= \Gamma\mathbf{C}(\mathbf{r},t) + \epsilon\mathbf{W}\boldsymbol{\eta}(\mathbf{r},t), \end{aligned} \tag{14}$$

where $\mathbf{W}$, $\mathbf{V}$ and $\boldsymbol{\eta}(\mathbf{r},t)$ are defined as above and $\Gamma = \epsilon\mathbf{W}\mathbf{V}(\boldsymbol{\alpha}_n(\mathbf{r}))$.

The equations (8) and (13), which combine the FDOT forward problem with the compartment model equations constitute the set of equations which will be used for the direct reconstruction of pharmacokinetic rates, volume fractions, and fluorescence concentrations in different compartments.

IV. DIRECT RECONSTRUCTION OF PHARMACOKINETIC RATE AND CONCENTRATION IMAGES FROM OPTICAL FLUX MEASUREMENTS

The source detector measurements in (9) are collected at discrete time instances, $t = kT$, $k = 0, 1, ...$, where T is the sampling period. Therefore, the continuous model described in (8) and (14) has to be discretized. To simplify our notation, we shall use $\mathbf{C}(\mathbf{r},k) = \mathbf{C}(\mathbf{r},kT)$ and $\mathbf{\Psi}(\mathbf{r},k) = \mathbf{\Psi}(\mathbf{r},kT)$.

Let $\boldsymbol{\theta}_n$ denote the discrete-time parameter vector of the pharmacokinetic rates and volume fractions. The parameter vector $\boldsymbol{\theta}_n(\mathbf{r})$ can be either time dependent or time independent. The formulation given in this work can be used for both cases. In our case, the pharmacokinetic rates and the volume fractions are time independent. However, in order to estimate $\boldsymbol{\theta}_n$ within the EKF framework, the following dynamic model is introduced:

$$\boldsymbol{\theta}_n(\mathbf{r},k+1) = \boldsymbol{\theta}_n(\mathbf{r},k) + \varsigma(\mathbf{r},k), \tag{15}$$

where $\varsigma(\mathbf{r},k)$ is a zero mean white noise process with covariance matrix $\mathbf{S}$. The details of the dynamic model introduced for the joint estimation of the system parameters and the states can be found in [3].

1) A priori information for Pharmacokinetic Rates and Volume Fractions: To improve the robustness of estimates of the parameters, we impose *a priori* information on pharmacokinetic parameters and volume fractions. Here, we assume that there is no information about the tumor structure and used a 4-pixel neighborhood model.

Using the 4-pixel neighbor model with equal weights β, the random process, $\boldsymbol{\theta}_n(\mathbf{r},k)$, can be modeled as:

$$\boldsymbol{\theta}_n(r_j,k+1) = \beta\boldsymbol{\theta}_n(r_j,k) + \beta\sum_{i=1}^{4}\boldsymbol{\theta}_n(r_{ji},k) + \varsigma(\mathbf{r},k). \tag{16}$$

2) Extended Kalman Filter (EKF) Formulation: In our state-space model, (8) and (9), both the states (concentrations) and system parameters (pharmacokinetic rates and volume fractions) are unknown. In this case, the state-space model can be regarded as a non-linear model in which system parameters and states are combined to form the new states of the non-linear model. This system is then linearized and solved for the unknown states using EKF framework [3]. To solve for the concentrations and the unknown parameters, the parameter vector $\boldsymbol{\theta}_n(\mathbf{r},k)$ is inserted into the concentration vector $\mathbf{C}(\mathbf{r},k)$ as:

$$\begin{bmatrix} \mathbf{C}(\mathbf{r},k+1) \\ \boldsymbol{\theta}_n(r_j,k+1) \end{bmatrix} = \begin{bmatrix} \mathbf{K}(\boldsymbol{\theta}_n(\mathbf{r},k))\mathbf{C}(\mathbf{r},k) \\ \beta\boldsymbol{\theta}_n(r_j,k) + \beta\sum_{i=1}^{4}\boldsymbol{\theta}_n(r_{ji},k) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\omega}(\mathbf{r},k) \\ \varsigma(\mathbf{r},k) \end{bmatrix}. \tag{17}$$

The measurement equation for the non-linear case is given as:

$$\mathbf{\Psi}(\mathbf{r},k) = \begin{bmatrix} g(\mathbf{V}(\boldsymbol{\alpha}_n(\mathbf{r}))\mathbf{C}(\mathbf{r},k)) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{C}(\mathbf{r},k) \\ \boldsymbol{\theta}_n(\mathbf{r},k) \end{bmatrix} + \eta(\mathbf{r},k). \tag{18}$$

Without explicit proof, the extended Kalman filtering algorithm for simultaneous estimation of concentrations and parameters for the nonlinear case, (13), is given follows:

$$\begin{bmatrix} \hat{\mathbf{C}}(\mathbf{r},0) \\ \hat{\boldsymbol{\theta}}_n(\mathbf{r},0) \end{bmatrix} = \begin{bmatrix} E(\mathbf{C}(\mathbf{r},0)) \\ \hat{\boldsymbol{\theta}}_n(\mathbf{r},0) \end{bmatrix}. \tag{19}$$

$$\mathbf{P}_{0,0} = \begin{bmatrix} Var(\mathbf{C}(\mathbf{r},0)) & 0 \\ 0 & \mathbf{S}_d \end{bmatrix}, \tag{20}$$

where E denotes the expected value of $\mathbf{C}(\mathbf{r},0)$, $\mathbf{P}$ is the error covariance matrix, and $\mathbf{S}_d$ is the preassigned covariance matrix of the unknown system parameters.

The following equations describe how the concentration estimates and error covariance matrix are updated at the kth time instant given all the measurements up to $(k-1)^{th}$ time instant. For $k = 1, 2, ...,$

$$\begin{bmatrix} \hat{\mathbf{C}}(\mathbf{r},k|k-1) \\ \hat{\boldsymbol{\theta}}_n(r_j,k|k-1) \end{bmatrix} = \begin{bmatrix} \mathbf{K}(\hat{\boldsymbol{\theta}}_n(r,k-1))\hat{\mathbf{C}}(\mathbf{r},k-1) \\ \beta\hat{\boldsymbol{\theta}}_n(r_j,k-1) + \beta\sum_{i=1}^{4}\hat{\boldsymbol{\theta}}_n(r_{ji},k) \end{bmatrix} \tag{21}$$

$$\mathbf{P}_{k|k-1} = \mathbf{J}(k-1)\mathbf{P}_{k-1,k-1}\mathbf{J}^T(k-1) + \begin{bmatrix} \mathbf{Q}_d & 0 \\ 0 & \mathbf{S}_d \end{bmatrix}, \tag{22}$$

where $\mathbf{J}$ is the Jacobian matrix given by

$$\mathbf{J}(k) = \begin{bmatrix} \mathbf{K}(\hat{\boldsymbol{\theta}}_n(\mathbf{r},k)) & \frac{\partial}{\partial\boldsymbol{\theta}}[\mathbf{K}(\hat{\boldsymbol{\theta}}_n(\mathbf{r},k))\hat{\mathbf{C}}(\mathbf{r},k)] \\ \frac{\partial}{\partial\mathbf{C}}\hat{\boldsymbol{\theta}}_n(\mathbf{r},k) & \frac{\partial}{\partial\hat{\boldsymbol{\theta}}_n}\hat{\boldsymbol{\theta}}_n(\mathbf{r},k) \end{bmatrix}. \tag{23}$$

$$\mathbf{G}_k = \mathbf{P}_{k,k-1}\mathbf{\Lambda}^T[\mathbf{\Lambda}\mathbf{P}_{k,k-1}\mathbf{\Lambda}^T + \mathbf{R}]^{-1}, \tag{24}$$

where $\mathbf{G}_k$ is the recursive Kalman gain, $\mathbf{R}$ is the covariance matrix of the measurements and $\mathbf{\Lambda}$ is:

$$\mathbf{\Lambda} = \begin{bmatrix} \frac{\partial}{\partial \mathbf{C}} g(\mathbf{V}(\boldsymbol{\theta}_n(\mathbf{r}, k|k-1)\mathbf{C}(\mathbf{r}, k|k-1)) \\ \cdot \\ \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{\partial}{\partial \mathbf{C}} g(\mathbf{V}(\boldsymbol{\theta}_n(\mathbf{r}, k|k-1)\mathbf{C}(\mathbf{r}, k|k-1\right) \end{bmatrix}^T, \tag{25}$$

$$\mathbf{P}_{k,k} = [\mathbf{I} - \mathbf{G}_k\mathbf{\Lambda}]\mathbf{P}_{k,k-1}. \tag{26}$$

where $\mathbf{I}$ is the identity matrix.

$$\begin{bmatrix} \hat{\mathbf{C}}(\mathbf{r}, k) \\ \hat{\boldsymbol{\theta}}_n(\mathbf{r}_j, k) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{C}}(\mathbf{r}, k|k-1) \\ \beta\hat{\boldsymbol{\theta}}_n(r_j, k|k-1) + \beta\sum_{i=1}^{4}\hat{\boldsymbol{\theta}}_n(r_{ji}, k|k-1) \end{bmatrix}$$

$$+\mathbf{G}_k(\mathbf{\Psi}(\mathbf{r}, k) - g(\mathbf{V}(\boldsymbol{\theta}_n(\mathbf{r}, k|k-1)\mathbf{C}(\mathbf{r}, k|k-1)).$$

V. SIMULATIONS AND RESULTS

To validate our approach, we performed a simulation study based on the two-compartment model proposed for ICG. Using physiologically correct values for pharmacokinetic rates, k_{in}, k_{out}, k_{elm}, and volume fractions, v_e, v_p, around the tumor region, a set of time series data, $\mathbf{\Psi}(\mathbf{r}, t)$, was generated from a simulated domain with tissue-like characteristics. To generate the synthetic measurements, the diffusion equation was solved numerically using FEM algorithm with Robin type boundary conditions [10]. The simulation used a modulation frequency 300 MHz. The phantom is 6cm by 6cm in size, and it is discretized into 24 by 24 voxels each of size 0.25cm by 0.25cm. The 24 sources and 24 detectors are arranged throughout the boundary sequentially. The maximum transition rates of k_{in} and k_{out} are simulated at the center of the image of size 1cm by 1cm and smoothly decreased towards the boundaries. Figures 2a and 3a display the true images of pharmacokinetic rates k_{in} and k_{out}. Figures 2b and 3b display the corresponding reconstructed images. We observe that there is a good agreement between the true and the estimated images in terms of localization of the heterogeneities. We calculated the percent error between the true images and the reconstructed images by the ratio of the L^2 norm of the error and the L^2 norm of the true image. The percent error for k_{in} and k_{out} are 11.95 % and 13.11 %, respectively.

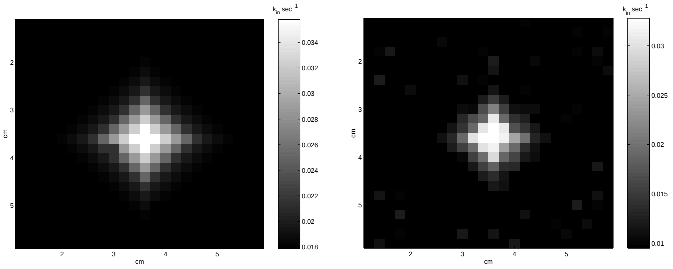


Fig. 2. Pharmacokinetic rate images of k_{in} : (a) true, and (b) reconstructed.

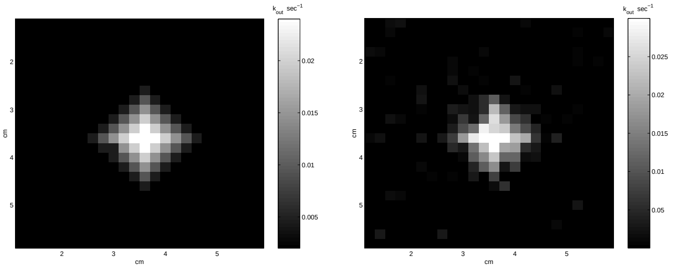


Fig. 3. Pharmacokinetic rate images of k_{out} : (a) true, and (b) reconstructed.

VI. CONCLUSION

In this paper, we provided a formulation for the reconstruction of spatially resolved pharmacokinetic rate images of fluorescence agents directly from the boundary photon flux measurements. We performed a simulation study using a numerical phantom. Reconstructed images with small errors show that the algorithm can be used for real data analysis. In the near feature, we plan to apply the proposed algorithm to the ICG concentration data acquired from breast tumors.

REFERENCES

[1] E. M. Sevick-Muraca et al., "Fluorescence and absorption contrast mechanisms for biomedical optical imaging using frequency-domain techniques," *Photochem. Photobiol.*, Vol.66, pp. 55, 1997.

[2] L. A. Bauer, *Applied Clinical Pharmacokinetics'*, McGraw-Hill, New York, 2001.

[3] B. Alacam, B. Yazici, X. Intes, B. Chance, " Extended Kalman Filtering for the Modeling and Analysis of ICG Pharmacokinetics in Cancerous Tumors using NIR Optical Methods", *IEEE Tran. in Biomedical Engineering*, in Press.

[4] D.J. Cuccia et al., "*In vivo* quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration," *Applied Optics*, vol. 42, no. 1, pp. 2940-2950, June 2003.

[5] X. Intes, J. Ripoll, Y. Chen, S. Nioka, A. G. Yodh, B. Chance, "In $vivo$ continuous-wave optical breast imaging enhanced with Indocyanine Green," *Med. Phys.*, vol. 30-6, pp.1039-1047, June 2003.

[6] C. Cobelli, D. Foster, and G. Toffolo, *Tracer Kinetics in Biomedical Research*, Kluwer Academic/Plenum, New York, c2000.

[7] M. Kamasak et al., "Direct reconstruction of kinetic parameter images from dynamic PET data," 37th *Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, pp. 19191923, 2003.

[8] A. B. Milstein, K. J. Webb, and C. A. Bouman, "Estimation of kinetic model parameters in fluorescence optical diffusion tomography," J. Opt. Soc. Am., Vol. 22, No. 7, pp. 1357-1368.

[9] B. Alacam et al., "Spatially Resolved Pharmacokinetic Rate Images of ICG using Near Infrared Optical Methods", Proc. of 2006 SPIE Photonic West, San Jose, CA, 21-26 January 2006, In Press.

[10] S. C. Brenner, and L. R. Scott, *The mathematical theory of finite element methods*, Springer, Berlin, 1994.

IOP Publishing Inverse Problems

Inverse Problems **23** (2007) 1115–1133 doi:10.1088/0266-5611/23/3/017

Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: I

Murat Guven[1], **Birsen Yazici**[1], **Kiwoon Kwon**[1], **Eldar Giladi**[2] **and Xavier Intes**[3]

[1] Electrical, Computer, and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY, USA
[2] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, USA
[3] Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA

E-mail: yazici@ecse.rpi.edu

Received 27 April 2006, in final form 31 January 2007
Published 15 May 2007
Online at stacks.iop.org/IP/23/1115

Abstract

In diffuse optical tomography (DOT), the discretization error in the numerical solutions of the forward and inverse problems results in error in the reconstructed optical images. In this first part of our work, we analyse the error in the reconstructed optical absorption images, resulting from the discretization of the forward and inverse problems. Our analysis identifies several factors which influence the extent to which the discretization impacts on the accuracy of the reconstructed images. For example, the mutual dependence of the forward and inverse problems, the number of sources and detectors, their configuration and their orientation with respect to optical absorptive heterogeneities, and the formulation of the inverse problem. As a result, our error analysis shows that the discretization of one problem cannot be considered independent of the other problem. While our analysis focuses specifically on the discretization error in DOT, the approach can be extended to quantify other error sources in DOT and other inverse parameter estimation problems.

1. Introduction

Imaging in diffuse optical tomography (DOT) comprises two interdependent stages which seek solutions to the forward and inverse problems. The forward problem is associated with describing the near-infrared (NIR) light propagation, while the objective of the inverse problem is to estimate the unknown optical parameters from boundary measurements [2].

There are a variety of factors that affect the accuracy of the DOT imaging, such as model mismatch (due to the light propagation model and/or linearization of the inverse problem), measurement noise, discretization, numerical algorithm efficiency and inverse problem formulation. In this two-part study, we focus on the effect of discretization of the

© 2007 IOP Publishing Ltd Printed in the UK

forward and inverse problems. In the first part of our work, we present an error analysis to show the effect of discretization on the accuracy of the reconstructed optical absorption images. We identify the factors specific to the imaging problem, which determine the extent to which the discretization impacts on the accuracy of the reconstructed optical absorption images. In the following, first, we use the error analysis to develop novel adaptive discretization algorithms for the forward and inverse problems to reduce the error in the reconstructed optical images resulting from discretization. Next, we present numerical experiments that support the main results of part I and demonstrate the effectiveness of the developed adaptive mesh generation algorithms.

There has been extensive research on the estimation of discretization error in the solutions of partial differential equations (PDEs) [1, 5–7, 21, 22]. In contrast, relatively little has been published in the area of parameter estimation problems governed by PDEs. See for example [8] for an *a posteriori* error estimate for the Lagrangian in the inverse scattering problem for the time-dependent acoustic wave equation and [19] for *a posteriori* error estimates for distributed elliptic optimal control problems. In the area of DOT, it was numerically shown that the approximation errors resulting from the discretization of the forward problem can lead to significant errors in the reconstructed optical images [3]. However, an analysis regarding the error in the reconstructed optical images resulting from discretization has not been reported so far.

In this work, we model the forward problem by the frequency-domain diffusion equation. For the inverse problem, we focus on the estimation of the absorption coefficient. We consider the linear integral equation resulting from the iterative linearization of the inverse problem based on Born approximation and use zeroth-order Tikhonov regularization to address the illposedness of the resulting integral equation. We use finite elements with first-order Lagrange basis functions to discretize the forward and inverse problems and analyse the effect of the discretization of each problem on the reconstructed optical absorption image. Our analysis describes the dependence of the image quality on the optical image properties, the configuration of the source and detectors, the orientation of the source and detectors with respect to absorptive heterogeneities, and on the regularization parameter in addition to the discretization error in the solution of each problem. In our analysis, we first consider the impact of the inverse problem discretization when there is no discretization error in the solution of the forward problem, and provide a bound for the resulting error in the reconstructed optical image. Next, we analyse the effect of the forward problem discretization on the accuracy of the reconstructed image without discretizing the inverse problem, and obtain another bound for the resulting error in the reconstructed optical image. We see that each error bound comprises the discretization error in the corresponding problem solution, scaled spatially by the solutions of both problems. This is a direct consequence of the fact that the inverse problem solution depends on the model defined by the forward problem. As a result, the error analysis yields specific error estimates which are different than the conventional discretization error estimates (see equations (3.8)–(3.9) and (4.14)) which only take into account the smoothness and support of the function of interest, and the finite-dimensional space of approximating functions [9]. We further discuss the use of other basis functions and methods in the discretization of the forward and inverse problems and explain how the error bounds can be modified accordingly. Finally, we extend our analysis to show the effect of noise on the accuracy of the reconstructed optical images. Our analysis shows that the presence of noise results in error terms in addition to the error in the reconstructed optical images induced by the discretization of the forward and inverse problems.

This work not only provides an insight into the error in reconstructed optical absorption images resulting from discretization, but also motivates the development of novel adaptive mesh generation algorithms to address this error [14]. In addition, the analysis presented in this work provides a means to identify and analyse the error in the reconstructed optical images resulting from the linearization of the Lippmann–Schwinger-type equations [10] using Born approximation [15]. Furthermore, the error analysis introduced in this paper is not limited to DOT, and can easily be extended for use in similar inverse parameter estimation problems such as electrical impedance tomography, bioluminescence tomography, optical fluorescence tomography, microwave imaging etc, in all of which the inverse problem can be interpreted in terms of a linear integral equation, whose kernel is the solution of a PDE that models the forward problem.

The outline of this paper is as follows: section 2 defines the forward and inverse problems. In section 3, we discuss the discretization of the forward and inverse problems. In section 4, we present two theorems that summarize the impact of discretization on the accuracy of the reconstructed optical images, which is followed by section 5. The appendices include results regarding the boundedness and compactness of the linear integral operator used to define the inverse problem, and the proof for the convergence of the inverse problem discretization.

2. Forward and inverse problems

In this section, we describe the model for NIR light propagation and define the forward and inverse DOT problems. Table 1 provides a list of the notation and table 2 provides the definition of function spaces and norms used throughout the paper. We note that we use calligraphic letters to denote the operators, e.g. $\mathcal{A}_a, \mathcal{I}, \mathcal{K}$ etc.

2.1. Forward problem

We use the following boundary value problem to model the NIR light propagation in a bounded domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$ [2, 9]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_j(\mathbf{x}) + \left(\mu_a(\mathbf{x}) + \frac{\mathrm{i}\omega}{c}\right) g_j(\mathbf{x}) = Q_j(\mathbf{x}) \qquad \mathbf{x} \in \Omega, \tag{2.1}$$

$$g_j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_j}{\partial n}(\mathbf{x}) = 0 \qquad \mathbf{x} \in \partial\Omega, \tag{2.2}$$

where $g_j(\mathbf{x})$ is the photon density at $\mathbf{x}$, Q_j is the point source located at $\mathbf{x}_s^j$, $j = 1, \ldots, N_s$, where N_s is the number of sources, $D(\mathbf{x})$ is the diffusion coefficient and $\mu_a(\mathbf{x})$ is the absorption coefficient at $\mathbf{x}$, $\mathrm{i} = \sqrt{-1}$, ω is the modulation frequency of the source, c is the speed of the light, $a = (1 + R)/(1 - R)$ where R is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial \cdot /\partial n$ denotes the directional derivative along the unit normal vector on the boundary. Note that we assume the diffusion coefficient is isotropic. For the general anisotropic material, see [17].

The adjoint problem [2] associated with (2.1)–(2.2) is given by the following boundary value problem:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^*(\mathbf{x}) + \left(\mu_a(\mathbf{x}) - \frac{\mathrm{i}\omega}{c}\right) g_i^*(\mathbf{x}) = 0 \qquad \mathbf{x} \in \Omega, \tag{2.3}$$

$$g_i^*(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_i^*}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \qquad \mathbf{x} \in \partial\Omega, \tag{2.4}$$

where Q_i^* is the adjoint source located at $\mathbf{x}_d^i$, $i = 1, \ldots, N_d$, where N_d is the number of detectors. We note that we approximate the point source Q_j in (2.1) and the adjoint source

Table 1. Definition of variables functions, and operators.

Notation	Explanation
Ω	Bounded domain in $\mathbb{R}^3$ with Lipschitz boundary
$\partial\Omega$	Lipschitz boundary of Ω
$\mathbf{x}$	Position vector in $\Omega \cup \partial\Omega$
$g_j(\mathbf{x})$	Solution of the diffusion equation at $\mathbf{x}$ for the jth point source located at $\mathbf{x}_s^j$
$g_i^*(\mathbf{x})$	Solution of the adjoint problem at $\mathbf{x}$ for the ith adjoint source located at $\mathbf{x}_d^i$
$G_j(\mathbf{x})$	Finite-element approximation of g_j at $\mathbf{x}$
$G_i^*(\mathbf{x})$	Finite-element approximation of g_i^* at $\mathbf{x}$
$e_j(\mathbf{x})$	The discretization error at $\mathbf{x}$ in the finite-element approximation of g_j
$e_i^*(\mathbf{x})$	The discretization error at $\mathbf{x}$ in the finite-element approximation of g_i^*
$\alpha(\mathbf{x})$	Small perturbation over the background μ_a at $\mathbf{x}$
$\Gamma_{i,j}$	Differential measurement at the ith detector due to the jth source
$\mathcal{A}_a$	The matrix-valued operator mapping $\alpha \in L^\infty(\Omega)$ to $\mathbb{C}^{N_d \times N_s}$
$\mathcal{A}_a^*$	The adjoint of $\mathcal{A}_a$ mapping from $\mathbb{C}^{N_d \times N_s}$ to $L^1(\Omega)$
$H_{i,j}(\mathbf{x})$	The kernel in $\mathcal{A}_a$ at $\mathbf{x}$ for the ith detector and the jth source
$H_{i,j}^*(\mathbf{x})$	The kernel in $\mathcal{A}_a^*$ at $\mathbf{x}$ for the ith detector and the jth source
$\gamma(\mathbf{x})$	$\mathcal{A}_a^*\Gamma$ at $\mathbf{x}$
λ	The regularization parameter
$\alpha^\lambda(\mathbf{x})$	Solution of the regularized inverse problem at $\mathbf{x}$
$\alpha_n^\lambda(\mathbf{x})$	Solution of the discretized regularized inverse problem with exact kernel at $\mathbf{x}$
$\tilde{\alpha}^\lambda(\mathbf{x})$	Solution of the regularized inverse problem with degenerate kernel at $\mathbf{x}$
$\tilde{\alpha}_n^\lambda(\mathbf{x})$	Solution of the discretized regularized inverse problem with degenerate kernel at $\mathbf{x}$

Table 2. Definition of function spaces and norms.

Notation	Explanation
$\overline{f}$	The complex conjugate of the function f
$C(\Omega)$	Space of continuous complex-valued functions on Ω
$C^k(\Omega)$	Space of complex-valued k-times continuously differentiable functions on Ω
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \mid ess\ \sup_\Omega \lvert f(\mathbf{x})\rvert < \infty\}$
$L^p(\Omega)$	$L^p(\Omega) = \{f \mid (\int_\Omega \lvert f(\mathbf{x})\rvert^p\,\mathrm{d}\mathbf{x})^{1/p} < \infty\}$, $p \in [1, \infty)$
$D_w^z f$	zth weak derivative of f
$H^p(\Omega)$	$H^p(\Omega) = \{f \mid (\sum_{\lvert z\rvert \leqslant p} \lVert D_w^z f\rVert_0^2)^{1/2} < \infty\}$, $p \in [1, \infty)$
$\lVert f\rVert_0$	The $L^2(\Omega)$ norm of f
$\lVert f\rVert_p$	The $H^p(\Omega)$ norm of f
$\lVert f\rVert_\infty$	The $L^\infty(\Omega)$ norm of f
$\lVert f\rVert_{L^p(\Omega)}$	The $L^p(\Omega)$ norm of f
$\lVert f\rVert_{0,m}$	The L^2 norm of f over the mth finite element Ω_m
$\lVert f\rVert_{p,m}$	The H^p norm of f over the mth finite element Ω_m

Q_i^* in (2.4) by Gaussian functions with sufficiently low variance, whose centres are located at $\mathbf{x}_s^j$ and $\mathbf{x}_d^i$, respectively.

In this work, we consider the finite-element approximations of the solutions of the forward problem. Hence, before we discretize the forward problem (see section 3.2), we consider the variational formulations of (2.1)–(2.2) and (2.3)–(2.4) by multiplying (2.1) by a test function $\overline{\phi} \in H^1(\Omega)$ and integrating over Ω [9]:

$$\int_\Omega \left[\nabla\overline{\phi}\cdot D\nabla g_j + \overline{\phi}\left(\mu_a + \frac{\mathrm{i}\omega}{c}\right) g_j - \overline{\phi} Q_j\right]\mathrm{d}\mathbf{x} + \frac{1}{2a}\int_{\partial\Omega} \overline{\phi} g_j\,\mathrm{d}l = 0, \tag{2.5}$$

where the boundary integral term results from the boundary condition (2.2).

Equivalently, we can express (2.5) by defining the sesquilinear form $b(\phi, g_j)$ [16]:

$$b(\phi, g_j) := A(\phi, g_j) + \left\langle \phi, \frac{1}{2a} g_j\right\rangle = (\phi, Q_j), \tag{2.6}$$

where

$$\begin{aligned}
A(\phi, g_j) &:= \int_\Omega \left[\nabla\overline{\phi}\cdot D\nabla g_j + \left(\mu_a + \frac{\mathrm{i}\omega}{c}\right)\overline{\phi} g_j\right]\mathrm{d}\mathbf{x},\\
(\phi, Q_j) &:= \int_\Omega \overline{\phi} Q_j\,\mathrm{d}\mathbf{x},\\
\left\langle \phi, \frac{1}{2a} g_j\right\rangle &:= \frac{1}{2a}\int_{\partial\Omega}\overline{\phi} g_j\,\mathrm{d}l.
\end{aligned}$$

Similarly, the variational problem for (2.3)–(2.4) can be formulated by defining the sesquilinear form $b^*(\phi, g_i^*)$:

$$b^*(\phi, g_i^*) := A(\phi, g_i^*) + \left\langle \phi, \frac{1}{2a} g_i^*\right\rangle = \left\langle \phi, \frac{1}{2a} Q_i^*\right\rangle, \tag{2.7}$$

where in $A(\phi, g_i^*)$, ω is replaced by $-\omega$.

The sesquilinear forms $b(\phi, g_j)$, $b^*(\phi, g_i^*)$ are continuous and positive definite for bounded D and μ_a [16]. As a result, the variational problems (2.6) and (2.7) have unique solutions, which follows from the Lax–Milgram lemma [9]. The solutions g_j and g_i^* of the variational problems (2.6) and (2.7) belong to $H^1(\Omega)$, which results from the H^1-boundedness of the Gaussian function that approximates the point source Q_j and the adjoint source Q_i^* [16]. Assuming $D, \mu_a \in C^1(\Omega)$ and noting that $Q_j, Q_i^* \in H^1(\Omega)$; the solutions g_j, g_i^* satisfy $g_j, g_i^* \in H^2_{loc}(\Omega)$ (in [12, chapter 6.3, theorem 2]). This last condition implies (in [12, chapter 5.6, theorem 6])

$$g_j, g_i^* \in C(\Omega). \tag{2.8}$$

2.2. Inverse problem

In this work, we focus on the estimation of the absorption coefficient; therefore, we assume $D(\mathbf{x})$ is known for all $\mathbf{x} \in \Omega \cup \partial\Omega$. To address the nonlinear nature of the inverse DOT problem, we consider an iterative algorithm based on repetitive linearization of the inverse problem using first-order Born approximation [2]. As a result, at each linearization step, the following linear integral equation relates the differential optical measurements to a small perturbation α on the absorption coefficient μ_a:

$$\Gamma_{i,j} = -\int_\Omega \overline{g_i^*(\mathbf{x})} g_j(\mathbf{x})\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{2.9}$$

$$\begin{aligned}
&:= \int_\Omega H_{i,j}(\mathbf{x})\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x}\\
&:= (\mathcal{A}_a\alpha)_{i,j},
\end{aligned} \tag{2.10}$$

where $H_{i,j} = -\overline{g_i^* g_j}$ is the kernel in the (i, j)th entry of the matrix-valued operator $\mathcal{A}_a : L^\infty(\Omega) \to \mathbb{C}^{N_d \times N_s}$, g_j is the solution of (2.6), g_i^* is the solution of (2.7), and $\Gamma_{i,j}$ is the (i, j)th entry in the vector $\Gamma \in \mathbb{C}^{N_d \times N_s}$, which represents the differential measurement at the ith detector due to the jth source. Note that approximating Q_i^* in (2.4) by a Gaussian function centred at $\mathbf{x}_d^i$ implies that $\Gamma_{i,j}$ corresponds to the scattered optical field evaluated at $\mathbf{x}_d^i$, after filtering it by that Gaussian function. Thus, the Gaussian approximation of the adjoint source models the finite size of the detectors. Similarly, approximating Q_j in (2.1) by a Gaussian function models the finite beam of the source.

The linear operator $\mathcal{A}_a : L^\infty(\Omega) \to \mathbb{C}^{N_d \times N_s}$ defined by (2.9) is compact and bounded by (see appendices A and B)

$$\|\mathcal{A}_a\|_{L^\infty(\Omega)\to l^1} \leqslant N_d N_s \max_i \|g_i^*\|_0 \max_j \|g_j\|_0. \tag{2.11}$$

For the given solution space $L^\infty(\Omega)$ for α, the compactness of the linear operator $\mathcal{A}_a$ implies the illposedness of (2.9). Hence, we regularize (2.9) with a zeroth-order Tikhonov regularization. This yields the following equation which defines our inverse problem at each linearization step:

$$\gamma := \mathcal{A}_a^* \Gamma = (\mathcal{A}_a^* \mathcal{A}_a + \lambda \mathcal{I})\alpha^\lambda \tag{2.12}$$

$$:= \mathcal{K}\alpha^\lambda, \tag{2.13}$$

where $\lambda > 0$ and α^λ is an approximation to α. In this representation, $\mathcal{I}$ is the identity operator and $\mathcal{A}_a^* : \mathbb{C}^{N_d \times N_s} \to L^1(\Omega)$ is the adjoint of $\mathcal{A}_a$, defined by

$$(\mathcal{A}_a^* \beta)(\mathbf{x}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x})\beta_{i,j} = \sum_{i,j}^{N_d,N_s} -g_i^*(\mathbf{x})\overline{g_j(\mathbf{x})}\beta_{i,j}, \tag{2.14}$$

for all $\beta \in \mathbb{C}^{N_d \times N_s}$, where $H_{i,j}^* := -g_i^* \overline{g_j}$ is the (i, j)th kernel in the adjoint operator $\mathcal{A}_a^*$. Let $\mathcal{A} := \mathcal{A}_a^* \mathcal{A}_a$, then $\mathcal{A} : L^\infty(\Omega) \to L^1(\Omega)$ is defined as follows:

$$(\mathcal{A}\alpha)(\mathbf{x}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x}) \int_\Omega H_{i,j}(\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})\,\mathrm{d}\acute{\mathbf{x}}$$

$$:= \int_\Omega \kappa(\mathbf{x}, \acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})\,\mathrm{d}\acute{\mathbf{x}}, \tag{2.15}$$

where $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ stands for the kernel of the integral operator $\mathcal{A}$ and is given by

$$\kappa(\mathbf{x}, \acute{\mathbf{x}}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x}) H_{i,j}(\acute{\mathbf{x}}). \tag{2.16}$$

Having defined the adjoint operator $\mathcal{A}_a^*$, we note that the operator $\mathcal{A} : L^\infty(\Omega) \to L^1(\Omega)$ is compact and that the operator $\mathcal{K} : L^\infty(\Omega) \to L^1(\Omega)$ is bounded by $\|\mathcal{K}\| \leqslant \|\mathcal{A}_a\|^2 + \lambda$. We assume that the solution $\alpha^\lambda \in L^\infty(\Omega)$ also satisfies $\alpha^\lambda \in H^1(\Omega)$. For the rest of the paper, we will denote $L^\infty(\Omega)$ and $L^1(\Omega)$ by X and Y, respectively.

3. Discretization of the inverse and forward problems

In this section, we outline the discretization of the inverse and forward problems.

3.1. Inverse problem discretization

In practice, we seek a finite-dimensional approximation to the solution of the inverse problem (2.13) at each linearization step. Therefore, we discretize (2.13) by projecting it onto a finite-dimensional subspace.

Let $X_n \subset X$ and $Y_n \subset Y$ denote a sequence of finite-dimensional subspaces of dimension $n = 1, 2, \ldots$, spanned by first-order Lagrange basis functions $\{L_1, \ldots, L_n\}$, and $\{\mathbf{x}_p\}$, $p = 1, \ldots, n$, be the set of collocation points on Ω. Then, the collocation method approximates the solution of (2.13) by an element $\alpha_n^\lambda \in X_n$ which satisfies

$$\left(\mathcal{K}\alpha_n^\lambda\right)(\mathbf{x}_p) = \gamma(\mathbf{x}_p), \qquad p = 1, \ldots, n, \tag{3.1}$$

where we express α_n^λ on a set $\{\Omega_m\}$ of finite elements, $m = 1, \ldots, N_\Delta$ such that $\bigcup_m^{N_\Delta} \Omega_m = \Omega$ as follows:

$$\alpha_n^\lambda(\mathbf{x}) = \sum_{k=1}^{n} a_k L_k(\mathbf{x}). \tag{3.2}$$

Note that in (3.2), $a_p = \alpha_n^\lambda(\mathbf{x}_p)$, $p = 1, \ldots, n$. Then, (3.1) can explicitly be written as

$$\lambda a_p + \sum_{k=1}^{n} a_k \int_\Omega \kappa(\mathbf{x}_p, \acute{\mathbf{x}}) L_k(\acute{\mathbf{x}})\,\mathrm{d}\acute{\mathbf{x}} = \gamma(\mathbf{x}_p), \qquad p = 1, \ldots, n. \tag{3.3}$$

Equivalently, the collocation method can be interpreted as a projection with the interpolation operator $\mathcal{P}_n : Y \to Y_n$ defined by [18]

$$\mathcal{P}_n f(\mathbf{x}) := \sum_{p=1}^{n} f(\mathbf{x}_p) L_p(\mathbf{x}), \qquad \mathbf{x} \in \Omega, \tag{3.4}$$

for all $f \in Y$. Then, (3.1) is equivalent to

$$\mathcal{P}_n \mathcal{K} \alpha_n^\lambda = \mathcal{P}_n \gamma. \tag{3.5}$$

3.2. Forward problem discretization

In this section, we consider the finite-element discretizations of (2.6) and (2.7), and use their solutions to approximate $H_{i,j}$ and $H_{i,j}^*$. As a result, we obtain finite-dimensional approximations to $\mathcal{K}$ and γ.

Let L_k denote the kth first-order Lagrange basis function. Replacing ϕ and g_j in (2.6) with their finite-dimensional counterparts $\Phi(\mathbf{x}) = \sum_{k=1}^{N_j} p_k L_k(\mathbf{x})$, $G_j(\mathbf{x}) = \sum_{k=1}^{N_j} c_k L_k(\mathbf{x})$, and replacing ϕ and g_i^* in (2.7) with $\Phi(\mathbf{x}) = \sum_{k=1}^{N_i} p_k L_k(\mathbf{x})$, $G_i^*(\mathbf{x}) = \sum_{k=1}^{N_i} d_k L_k(\mathbf{x})$ yields the matrix equations:

$$\mathbf{S}\mathbf{c}_j = \mathbf{q}_j, \tag{3.6}$$

$$\mathbf{S}^*\mathbf{d}_i = \mathbf{q}_i^*, \tag{3.7}$$

for $\mathbf{c}_j := \left[c_1, c_2, \ldots, c_{N_j}\right]^T$ and $\mathbf{d}_i := \left[d_1, d_2, \ldots, d_{N_i}\right]^T$. Here $\mathbf{S}$ and $\mathbf{S}^*$ are the finite-element matrices and $\mathbf{q}_j$ and $\mathbf{q}_i^*$ are the load vectors resulting from the finite-element discretization of (2.6) and (2.7). Note that for each source (detector), the dimension of the finite-element solution G_j (G_i^*) can be different; therefore, N_j (N_i) may vary.

The $H^1(\Omega)$ boundedness of the solutions g_j and g_i^* implies that the discretization errors e_j and e_i^* in G_j and G_i^* are bounded. Let $\left\{\Omega_m^j\right\}$ denote the set of linear elements used to discretize (2.6) for $m = 1, \ldots, N_\Delta^j$, such that $\bigcup_m^{N_\Delta^j} \Omega_m^j = \Omega$ for all $j = 1, \ldots, N_s$. Similarly, let $\left\{\Omega_n^i\right\}$ denote the set of linear elements used to discretize (2.7) for $n = 1, \ldots, N_\Delta^{*i}$, such that $\bigcup_n^{N_\Delta^{*i}} \Omega_n^i = \Omega$ for all $i = 1, \ldots, N_d$. Then, a bound for e_j and e_i^* on each finite element can be found by using the discretization error estimates (in [9, theorem 4.4.4]):

$$\|e_j\|_{0,m^j} \leqslant C\|g_j\|_{1,m^j} h_m^j, \tag{3.8}$$

$$\|e_i^*\|_{0,n^i} \leqslant C\|g_i^*\|_{1,n^i} h_n^i, \tag{3.9}$$

where C is a positive constant, $\|\cdot\|_{0,m^j}$ $(\|\cdot\|_{0,n^i})$ and $\|\cdot\|_{1,m^j}$ $(\|\cdot\|_{1,n^i})$ are respectively the L^2 and H^1 norms on $\Omega_m^j\left(\Omega_n^i\right)$, and h_m^j $\left(h_n^i\right)$ is the diameter of the smallest ball containing the finite element Ω_m^j $\left(\Omega_n^i\right)$ in the solution G_j (G_i^*).

3.3. Discretization of the inverse problem with operator approximations

Substituting the finite-element approximations G_j and G_i^* in (2.15) and (2.14), and using the resulting finite-dimensional operator approximations in (3.5), we obtain the following linear system in terms of $\tilde{\alpha}_n^\lambda$ which approximates α^λ:

$$\mathcal{P}_n\tilde{\mathcal{K}}\tilde{\alpha}_n^\lambda = \mathcal{P}_n\tilde{\gamma}. \tag{3.10}$$

In (3.10), the operator $\tilde{\mathcal{K}} : X \to Y$ is the finite-dimensional approximation of $\mathcal{K}$ in (2.13) and $\mathcal{P}_n\tilde{\mathcal{K}} : X_n \to Y_n$. Similarly,

$$\tilde{\gamma} := \tilde{\mathcal{A}}_a^*\Gamma, \tag{3.11}$$

where $\tilde{\mathcal{A}}_a^*$ is the approximation to the adjoint operator $\mathcal{A}_a^*$, obtained by substituting G_j and G_i^* in (2.14).

4. Discretization-based error analysis

As a result of the discretization of the forward and inverse problems, the reconstructed image $\tilde{\alpha}_n^\lambda$ in (3.10) is an approximation to the actual image α^λ. Thus, the accuracy of the reconstructed image depends on the error incurred by the discretization of the forward and inverse problems.

In this section, we analyse the effect of the discretization of the forward and inverse problems on the accuracy of DOT imaging. The analysis is carried out based on the inverse problem at each linearization defined by (2.13) and the associated kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$.

In this work, we follow an approach which allows us to separately analyse the effect of the discretization of each problem on the accuracy of the reconstructed optical image. In this respect, we first consider the impact of projection (i.e. inverse problem discretization) by the collocation method when the associated kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ in (2.13) is exact. Next, we explore the case in which the kernel is replaced by its finite-dimensional approximation (i.e. degenerate kernel) and analyse the effect of the forward problem discretization on the accuracy of the reconstructed image without projecting (2.13).

Our analysis reveals that even if the kernel is exact, the accuracy of the solution approximation α_n^λ in (3.5) resulting from the inverse problem discretization depends on the kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ of the integral operator. Likewise, the error in the reconstructed optical image due to the discretization of the forward problem is a function of the inverse problem solution. These results suggest that the discretization of the inverse and forward problems cannot be considered independent of each other.

4.1. Case 1. The kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ is exact

In this section, we show the effect of projection on the optical imaging accuracy. In the analysis, we assume that the kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ is exact. We first prove the convergence of the projection method for the operator $\mathcal{K}$, and then analyse the effect of projection on the imaging accuracy.

Clearly, the inverse operator $\mathcal{K}^{-1} : Y \to X$ exists since $\mathcal{K}$ is positive definite for $\lambda > 0$. Furthermore, by the compactness of $\mathcal{A}$ and Riesz theorem, the inverse operator $\mathcal{K}^{-1}$ is bounded by

$$\|\mathcal{K}^{-1}\|_{Y\to X} \leqslant \frac{1}{\lambda}. \tag{4.1}$$

Lemma. *Projection by the collocation method for the operator $\mathcal{K} : X \to Y$ converges. Specifically, the sequence of finite-dimensional operators $\mathcal{P}_n\mathcal{K} : X_n \to Y_n$ is invertible for sufficiently large n, and $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\alpha^\lambda \to \alpha^\lambda$, $n \to \infty$. Furthermore,*

$$\|(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\|_{X\to X_n} \leqslant C_M \frac{\|\mathcal{K}\|_{X\to Y}}{\lambda} \tag{4.2}$$

for some $C_M > 0$ independent of n.

Proof. See appendix C. □

Based on the lemma, the following theorem provides an upper bound for the $L^1(\Omega)$ norm of the error between the solution α^λ of (2.13) and the solution α_n^λ of (3.5).

Theorem 1. *Let $\{\Omega_m\}$ denote a set of linear finite elements used in the discretization of the inverse problem (2.13) for $m = 1, \ldots, N_\Delta$, such that $\bigcup_m^{N_\Delta} \Omega_m = \Omega$, and h_m be the diameter of the smallest ball that contains the* mth *element. Then,*

$$\begin{aligned}\left\|\alpha^\lambda - \alpha_n^\lambda\right\|_{L^1(\Omega)} \leqslant{}& C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n} \sum_{m=1}^{N_\Delta} \|\alpha^\lambda\|_{1,m} h_m \\ &+ \frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n} \max_{i,j}\|g_i^* g_j\|_{L^1(\Omega)} \sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m}\|\alpha^\lambda\|_{1,m} h_m,\end{aligned} \tag{4.3}$$

where C is a positive constant, V_Ω is the volume of Ω and $\mathcal{T}_n : Y \to X_n$ is a uniformly bounded operator given by $\mathcal{T}_n = \left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)^{-1}\mathcal{P}_n$.

Proof.

$$\begin{aligned}\alpha^\lambda - \alpha_n^\lambda &= [\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}]\alpha^\lambda \\ &= [\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}](\alpha^\lambda - \psi)\end{aligned} \tag{4.4}$$

since $[\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}]\psi = 0$, where $\psi \in X_n$ is the interpolant of α^λ [9]. Using (C.2),

$$\begin{aligned}[\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}] &= \mathcal{I} - \left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)^{-1}\frac{1}{\lambda}\mathcal{P}_n\mathcal{K} \\ &= \mathcal{I} - \mathcal{T}_n\frac{1}{\lambda}\mathcal{K},\end{aligned} \tag{4.5}$$

where $\mathcal{T}_n := \left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)^{-1}\mathcal{P}_n$ is a uniformly bounded operator (see appendix C). We use $\mathcal{K}$ defined by (2.13) and (4.5) in (4.4) to obtain

$$\alpha^\lambda - \alpha_n^\lambda = (\mathcal{I} - \mathcal{T}_n)(\alpha^\lambda - \psi) - \frac{\mathcal{T}_n}{\lambda}\mathcal{A}(\alpha^\lambda - \psi). \tag{4.6}$$

Then, we use the definition of $\mathcal{A}$ in (4.6) and find

$$\alpha^\lambda - \alpha_n^\lambda = (\mathcal{I} - \mathcal{T}_n)(\alpha^\lambda - \psi) - \frac{\mathcal{T}_n}{\lambda}\int_\Omega \kappa(\cdot, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})\,\mathrm{d}\acute{\mathbf{x}}. \tag{4.7}$$

This leads to

$$\begin{aligned}\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)} &\leqslant \|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n}\|\alpha^\lambda - \psi\|_{L^1(\Omega)} \\ &\quad + \frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\left\|\int_\Omega \kappa(\cdot, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})\,\mathrm{d}\acute{\mathbf{x}}\right\|_{L^1(\Omega)} \\ &\leqslant \|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n}\|\alpha^\lambda - \psi\|_{L^1(\Omega)} \\ &\quad + \frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\int_\Omega \mathrm{d}\mathbf{x}\int_\Omega |\kappa(\mathbf{x}, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}},\end{aligned} \tag{4.8}$$

The second term in (4.8) can be rewritten as

$$\begin{aligned}&\frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\int_\Omega \mathrm{d}\mathbf{x}\int_\Omega |\kappa(\mathbf{x}, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}} \\ &\quad = \frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\int_\Omega \mathrm{d}\mathbf{x}\left(\sum_{m=1}^{N_\Delta}\int_{\Omega_m} |\kappa(\mathbf{x}, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}}\right).\end{aligned} \tag{4.9}$$

Let e_α be the interpolation error:

$$e_\alpha := \alpha^\lambda - \psi. \tag{4.10}$$

Then, using (2.16),

$$\sum_{m=1}^{N_\Delta}\int_{\Omega_m} |\kappa(\mathbf{x}, \acute{\mathbf{x}})e_\alpha(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}} = \sum_{m=1}^{N_\Delta}\int_{\Omega_m}\left|\sum_{i,j}^{N_d,N_s} g_i^*(\mathbf{x})g_j(\mathbf{x})\overline{g_i^*(\acute{\mathbf{x}})}g_j(\acute{\mathbf{x}})e_\alpha(\acute{\mathbf{x}})\right|\mathrm{d}\acute{\mathbf{x}} \tag{4.11}$$

$$\begin{aligned}&\leqslant \sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s} |g_i^*(\mathbf{x})g_j(\mathbf{x})|\int_{\Omega_m} |g_i^*(\acute{\mathbf{x}})g_j(\acute{\mathbf{x}})||e_\alpha(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}} \\ &\leqslant \sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s} |g_i^*(\mathbf{x})g_j(\mathbf{x})|\|g_i^*g_j\|_{0,m}\|e_\alpha\|_{0,m},\end{aligned} \tag{4.12}$$

where (4.12) follows from the Schwarz' inequality. Note that $g_i^*g_j \in L^2(\Omega)$ by considering (2.8) holds up to the boundary $\partial\Omega$ (see [11, theorem 2.1]).

We now use (4.9) and (4.12) to obtain

$$\begin{aligned}&\frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\int_\Omega \mathrm{d}\mathbf{x}\left(\int_\Omega |\kappa(\mathbf{x}, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}}\right) \\ &\quad\leqslant \frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\int_\Omega \mathrm{d}\mathbf{x}\sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s} |g_i^*(\mathbf{x})g_j(\mathbf{x})|\|g_i^*g_j\|_{0,m}\|e_\alpha\|_{0,m}.\end{aligned} \tag{4.13}$$

Using the bound (4.13) in (4.8) and substituting the interpolation error bound [9]

$$\|e_\alpha\|_{0,m} \leqslant C\|\alpha^\lambda\|_{1,m}h_m, \tag{4.14}$$

and noting $\|e_\alpha\|_{L^1(\Omega)} \leqslant \sqrt{V_\Omega}\sum_{m=1}^{N_\Delta}\|e_\alpha\|_{0,m}$, we obtain

$$\begin{aligned}
\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)} \leqslant{}& C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\|\alpha^\lambda\|_{1,m}h_m \\
&+ \frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s}\|g_i^* g_j\|_{L^1(\Omega)}\|g_i^* g_j\|_{0,m}\|\alpha^\lambda\|_{1,m}h_m. \\
\leqslant{}& C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\|\alpha^\lambda\|_{1,m}h_m \\
&+ \frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\max_{i,j}\|g_i^* g_j\|_{L^1(\Omega)}\sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s}\|g_i^* g_j\|_{0,m}\|\alpha^\lambda\|_{1,m}h_m.
\end{aligned} \tag{4.15}$$

□

Remark 1.

(i) Theorem 1 shows the spatial dependence of the inverse problem discretization on the forward problem solution.

(ii) The first term in (4.15) suggests that the mesh of the inverse problem be refined where $\|\alpha^\lambda\|_1$ is large.

(iii) The second term in (4.15) shows that the term $\|\alpha^\lambda\|_{1,m}$ is scaled spatially by $\|g_i^* g_j\|_{0,m}$. Thus, the effect of the interpolation error e_α (see equation (4.10)) in the inverse problem solution is scaled spatially by the solution of the forward problem. As a result, the orientation of the sources and detectors with respect to the support of the optical heterogeneity determines the extent of the bound on $\left\|\alpha^\lambda - \alpha_n^\lambda\right\|_{L^1(\Omega)}$.

(iv) The regularization parameter affects the bound on $\left\|\alpha^\lambda - \alpha_n^\lambda\right\|_{L^1(\Omega)}$.

(v) Increasing the number of sources and detectors increases the bound on $\left\|\alpha^\lambda - \alpha_n^\lambda\right\|_{L^1(\Omega)}$.

Remark 2.

(i) Note that the conventional interpolation error estimate given in (4.14) depends only on the smoothness and support of α^λ, and the finite-dimensional space of approximating functions [9]. On the other hand, the error estimate (4.3) in theorem 1 shows that the accuracy of the reconstructed image α_n^λ depends on the orientation of the absorptive heterogeneity with respect to the sources and detectors, as well as on the bound (4.14) on the interpolation error.

(ii) An error bound similar to (4.3) follows if one uses the Galerkin method [18] instead of the collocation method for projection.

(iii) The interpolation error bound (4.14) can be modified based on the choice of the basis function in (3.2) and the smoothness of the solution α^λ (theorem 4.4.4. in [9]). For instance, if $\alpha^\lambda \in H^2(\Omega)$ and quadratic Lagrange basis functions are used, then (4.14) can be replaced by

$$\|e_\alpha\|_{0,m} \leqslant C\|\alpha^\lambda\|_{2,m}h_m^2,$$

for some $C > 0$.

(iv) An error bound similar to (4.3) can be derived for the error that occurs as a result of the discretization of the inverse problem in electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography and microwave imaging. Note that in all these imaging modalities, the forward problem is modelled by a PDE and the inverse problem can be interpreted in terms of a linear integral equation, whose kernel is related to the solution of this PDE.

(v) Let γ^δ be the perturbed right-hand side γ of (3.5) due to the presence of noise, such that $\|\gamma^\delta - \gamma\|_{L^1(\Omega)} \leqslant \delta$. Then, an additional term is introduced to the error bound in (4.3) due to this perturbation:

$$\begin{aligned}\left\|\alpha^\lambda - \alpha_n^\lambda\right\|_{L^1(\Omega)} \leqslant{}& C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n} \sum_{m=1}^{N_\Delta} \|\alpha^\lambda\|_{1,m} h_m \\ &+ \frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n} \max_{i,j}\|g_i^* g_j\|_{L^1(\Omega)} \sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m}\|\alpha^\lambda\|_{1,m} h_m \\ &+ \frac{C_M}{\lambda}\delta, \end{aligned} \tag{4.16}$$

where $C_M > 0$ is the constant in (4.2) with the use of first-order Lagrange basis functions (see appendix C). The additional term $C_M\delta/\lambda$ indicates that the choice of basis functions may be critical in the presence of noise.

4.2. Case 2. The kernel is degenerate

In this section, we first derive approximate upper bounds for the approximation errors $\|\tilde{\mathcal{K}} - \mathcal{K}\|$ and $\|\tilde{\gamma} - \gamma\|$, which result from the discretization of the forward problem. Then, we show the effect of these approximation errors on the accuracy of the reconstructed optical image. For notational convenience, we will drop the subscripts on the norms $\|\cdot\|$ where necessary.

The operator $\mathcal{K} : X \to Y$ is bounded with a bounded inverse $\mathcal{K}^{-1} : Y \to X$. By the finite-element approximation of the associated kernel, the sequence of bounded linear finite-dimensional operators $\tilde{\mathcal{K}}$ is norm convergent $\|\tilde{\mathcal{K}} - \mathcal{K}\| \to 0$; $N_j, N_i \to \infty$, for $j = 1, \ldots, N_s$ and $i = 1, \ldots, N_d$, and

$$\|\tilde{\mathcal{K}}^{-1}\|_{Y\to X} < 1/\lambda, \tag{4.17}$$

which can be obtained analogous to (4.1).

In the following, we derive an explicit approximation to the error $\|\tilde{\mathcal{K}} - \mathcal{K}\|$ in terms of the associated kernel and the discretization error in the kernel approximation. The result is then used to compute the error in the reconstructed optical image due to $\|\tilde{\mathcal{K}} - \mathcal{K}\|$.

By definition,

$$\|(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\alpha\|_{l^1} = \sum_{i,j}^{N_d,N_s} \left| \int_\Omega (\overline{g_i^*(\mathbf{x})} g_j(\mathbf{x}) - \overline{G_i^*(\mathbf{x})} G_j(\mathbf{x}))\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x} \right|, \tag{4.18}$$

where G_i^*, G_j are the finite-element approximations to g_i^* and g_j, respectively. We can expand $\overline{g_i^*} g_j - \overline{G_i^*} G_j$ as

$$\overline{g_i^*} g_j - \overline{G_i^*} G_j = \overline{e_i^*} e_j + G_j \overline{e_i^*} + \overline{G_i^*} e_j, \tag{4.19}$$

where $e_i^* := g_i^* - G_i^*$ and $e_j := g_j - G_j$. Replacing G_i^* and G_j respectively with $g_i^* - e_i^*$ and $g_j - e_j$, we get

$$\begin{aligned}\overline{g_i^*} g_j - \overline{G_i^*} G_j &= g_j \overline{e_i^*} + \overline{g_i^*} e_j - \overline{e_i^*} e_j \\ &\approx g_j \overline{e_i^*} + \overline{g_i^*} e_j, \end{aligned} \tag{4.20}$$

where we neglect the term $\overline{e_i^*} e_j$.

We can express $\mathcal{K} - \tilde{\mathcal{K}}$ as

$$\mathcal{K} - \tilde{\mathcal{K}} = \mathcal{A}_a^* \mathcal{A}_a - \tilde{\mathcal{A}}_a^* \tilde{\mathcal{A}}_a. \tag{4.21}$$

Following a similar approach as above,

$$\mathcal{A}_a^*\mathcal{A}_a - \tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a = (\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + \tilde{\mathcal{A}}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + (\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)\tilde{\mathcal{A}}_a. \tag{4.22}$$

As a result, the following condition holds

$$\|\tilde{\mathcal{K}} - \mathcal{K}\| \leqslant \|(\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\| + \|\tilde{\mathcal{A}}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + (\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)\tilde{\mathcal{A}}_a\|. \tag{4.23}$$

Since $\tilde{\mathcal{A}}_a = -(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + \mathcal{A}_a$, (4.23) can be rewritten as

$$\begin{aligned}\|\tilde{\mathcal{K}} - \mathcal{K}\| &= \|\mathcal{A}_a^*\mathcal{A}_a - \tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a\| \\ &\leqslant \|(\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\| + 2\|\mathcal{A}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\| \\ &\approx 2\|\mathcal{A}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\|,\end{aligned} \tag{4.24}$$

where we neglect the term $\|(\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\|$.

Similarly, $\|\tilde{\gamma} - \gamma\|$ can be interpreted as

$$\begin{aligned}\|\tilde{\gamma} - \gamma\|_{L^1(\Omega)} &= \int_\Omega \left|\sum_{i,j}^{N_d,N_s} (g_i^*(\mathbf{x})\overline{g_j(\mathbf{x})} - G_i^*(\mathbf{x})\overline{G_j(\mathbf{x})})\Gamma_{i,j}\right| \mathbf{dx} \\ &\approx \int_\Omega \left|\sum_{i,j}^{N_d,N_s} (e_i^*(\mathbf{x})\overline{g_j(\mathbf{x})} + g_i^*(\mathbf{x})\overline{e_j(\mathbf{x})})\Gamma_{i,j}\right| \mathbf{dx},\end{aligned} \tag{4.25}$$

where the error in $\Gamma_{i,j}$ due to discretization is neglected and the last approximation is derived similar to (4.20).

We now analyse the effect of the forward problem discretization on the accuracy of the reconstructed optical image. Let $\tilde{\alpha}^\lambda$ be the solution of

$$\tilde{\mathcal{K}}\tilde{\alpha}^\lambda = \tilde{\gamma}, \tag{4.26}$$

where $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ are the finite-dimensional approximations to $\mathcal{K}$ and γ, respectively. Then, by theorem 10.1 in [18], the error in the solution $\tilde{\alpha}^\lambda$ with respect to the actual solution α^λ is bounded by

$$\|\alpha^\lambda - \tilde{\alpha}^\lambda\| \leqslant \frac{1}{\lambda}\{\|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\| + \|\tilde{\gamma} - \gamma\|\}. \tag{4.27}$$

In the next theorem, we will expand the terms in (4.27) to show explicitly the effect of the forward problem discretization on the accuracy of the inverse problem solution.

Theorem 2. *Let $\{\Omega_m^j\}$ denote the set of linear elements used to discretize (2.6) for $m = 1, \ldots, N_\Delta^j$, such that $\bigcup_m^{N_\Delta^j} \Omega_m^j = \Omega$ and h_m^j be the diameter of the smallest ball that contains the element Ω_m^j in the solution G_j, for all $j = 1, \ldots, N_s$. Similarly, let $\{\Omega_n^i\}$ denote the set of linear elements used to discretize (2.7) for $n = 1, \ldots, N_\Delta^{*i}$, such that $\bigcup_n^{N_\Delta^{*i}} \Omega_n^i = \Omega$ and h_n^i be the diameter of the smallest ball that contains the element Ω_n^i in the solution G_i^*, for all $i = 1, \ldots, N_d$. Then, a bound for the error between the solution α^λ of (2.13) and the solution $\tilde{\alpha}^\lambda$ of (4.26) due to the approximations $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ is given by*

$$\begin{aligned}\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)} \leqslant \frac{C}{\lambda}\max_{i,j}\|g_i^* g_j\|_{L^1(\Omega)} \Bigg(&\sum_{i=1}^{N_d}\sum_{n,j}^{N_\Delta^{*i},N_s} \left(2\|g_j\alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty\|g_j\|_{0,n^i}\right)\|g_i^*\|_{1,n^i} h_n^i \\ &+ \sum_{j=1}^{N_s}\sum_{m,i}^{N_\Delta^j,N_d} (2\|g_i^*\alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty\|g_i^*\|_{0,m^j})\|g_j\|_{1,m^j} h_m^j\Bigg),\end{aligned} \tag{4.28}$$

where C is a positive constant.

Proof. Using (4.24), (4.18) and (4.20), we can write

$$\begin{aligned}
\|(\tilde{\mathcal{K}}-\mathcal{K})\alpha^{\lambda}\|_{L^1(\Omega)} &\approx 2\|\mathcal{A}_a^*(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\alpha^{\lambda}\|_{L^1(\Omega)}\\
&\approx 2\left\|\sum_{i,j}^{N_d,N_s} g_i^*(\cdot)\overline{g_j(\cdot)}\int_\Omega (g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})}+\overline{g_i^*(\acute{\mathbf{x}})}e_j(\acute{\mathbf{x}}))\alpha^{\lambda}(\acute{\mathbf{x}})\,\mathrm{d}\acute{\mathbf{x}}\right\|_{L^1(\Omega)}\\
&\leqslant 2\max_{i,j}\|g_i^*g_j\|_{L^1(\Omega)}\sum_{i,j}^{N_d,N_s}\int_\Omega |(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})}+\overline{g_i^*(\acute{\mathbf{x}})}e_j(\acute{\mathbf{x}}))\alpha^{\lambda}(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}}. \qquad (4.29)
\end{aligned}$$

An upper bound for the integral in (4.29) can be obtained as follows:

$$\begin{aligned}
&\int_\Omega |(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})}+\overline{g_i^*(\acute{\mathbf{x}})}e_j(\acute{\mathbf{x}}))\alpha^{\lambda}(\acute{\mathbf{x}})|\,\mathrm{d}\acute{\mathbf{x}}\\
&\qquad\leqslant \sum_{n=1}^{N_\Delta^{*i}}\|e_i^*\|_{0,n^i}\|g_j\alpha^{\lambda}\|_{0,n^i}+\sum_{m=1}^{N_\Delta^{j}}\|e_j\|_{0,m^j}\|g_i^*\alpha^{\lambda}\|_{0,m^j}. \qquad (4.30)
\end{aligned}$$

Note that $g_j\alpha^{\lambda}\in L^2(\Omega)$ since $|g_j\alpha^{\lambda}|\leqslant|g_j|\|\alpha^{\lambda}\|_\infty$. Similarly, $g_i^*\alpha^{\lambda}\in L^2(\Omega)$ since $|g_i^*\alpha^{\lambda}|\leqslant|g_i^*|\|\alpha^{\lambda}\|_\infty$. Using (4.30) in (4.29),

$$\begin{aligned}
\|(\tilde{\mathcal{K}}-\mathcal{K})\alpha^{\lambda}\|_{L^1(\Omega)} &\leqslant 2\max_{i,j}\|g_i^*g_j\|_{L^1(\Omega)}\\
&\quad\times\left(\sum_{i=1}^{N_d}\sum_{n,j}^{N_\Delta^{*i},N_s}\|e_i^*\|_{0,n^i}\|g_j\alpha^{\lambda}\|_{0,n^i}+\sum_{j=1}^{N_s}\sum_{m,i}^{N_\Delta^{j},N_d}\|e_j\|_{0,m^j}\|g_i^*\alpha^{\lambda}\|_{0,m^j}\right).
\end{aligned} \qquad (4.31)$$

To compute an upper bound for $\|\tilde{\gamma}-\gamma\|$ using (4.25), we first write

$$\begin{aligned}
&\int_\Omega\left|\sum_{i,j}^{N_d,N_s}(e_i^*(\mathbf{x})\overline{g_j(\mathbf{x})}+g_i^*(\mathbf{x})\overline{e_j(\mathbf{x})})\Gamma_{i,j}\right|\mathrm{d}\mathbf{x}\\
&\qquad\leqslant\max_{i,j}|\Gamma_{i,j}|\int_\Omega\sum_{i,j}^{N_d,N_s}|e_i^*(\mathbf{x})\overline{g_j(\mathbf{x})}+g_i^*(\mathbf{x})\overline{e_j(\mathbf{x})}|\,\mathrm{d}\mathbf{x}\\
&\qquad\leqslant\max_{i,j}|\Gamma_{i,j}|\left(\sum_{i=1}^{N_d}\sum_{n,j}^{N_\Delta^{*i},N_s}\|e_i^*\|_{0,n^i}\|g_j\|_{0,n^i}+\sum_{i=j}^{N_s}\sum_{m,i}^{N_\Delta^{j},N_d}\|g_i^*\|_{0,m^j}\|e_j\|_{0,m^j}\right). \qquad (4.32)
\end{aligned}$$

Noting (2.9),

$$\max_{i,j}|\Gamma_{i,j}|\leqslant\max_{i,j}\|g_i^*g_j\|_{L^1(\Omega)}\|\alpha\|_\infty, \qquad (4.33)$$

which leads to

$$\begin{aligned}
&\max_{i,j}|\Gamma_{i,j}|\left(\sum_{i=1}^{N_d}\sum_{n,j}^{N_\Delta^{*i},N_s}\|e_i^*\|_{0,n^i}\|g_j\|_{0,n^i}+\sum_{i=j}^{N_s}\sum_{m,i}^{N_\Delta^{j},N_d}\|g_i^*\|_{0,m^j}\|e_j\|_{0,m^j}\right)\\
&\quad\leqslant\max_{i,j}\|g_i^*g_j\|_{L^1(\Omega)}\|\alpha\|_\infty\left(\sum_{i=1}^{N_d}\sum_{n,j}^{N_\Delta^{*i},N_s}\|e_i^*\|_{0,n^i}\|g_j\|_{0,n^i}+\sum_{i=j}^{N_s}\sum_{m,i}^{N_\Delta^{j},N_d}\|g_i^*\|_{0,m^j}\|e_j\|_{0,m^j}\right).
\end{aligned} \qquad (4.34)$$

We now use (4.31), (4.34), the corresponding discretization error estimates (3.8)–(3.9), and (4.27) to obtain (4.28). ☐

Remark 3.

(i) Theorem 2 suggests the use of meshes designed individually for the solutions G_j, $j = 1, \ldots, N_s$ and G_i^*, $i = 1, \ldots, N_d$.

(ii) Theorem 2 states explicitly the effect of the forward problem discretization on the accuracy of the inverse problem solution. In this context, theorem 2 suggests a discretization scheme for the forward problem, where the discretization criterion is based on the inverse problem solution accuracy, rather than the accuracy of the forward problem solution.

(iii) For each source, when solving for G_j, h_m^j has to be kept small where $(2\|g_i^*\alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty\|g_i^*\|_{0,m^j})\|g_j\|_{1,m^j}$ is large. Note that $\|g_j\|_{1,m^j}$ will be large on the elements close to the jth source.

(iv) For each detector, when solving for G_i^*, h_n^i has to be kept small where $(2\|g_j\alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty\|g_j\|_{0,n^i})\|g_i^*\|_{1,n^i}$ is large. Note that $\|g_i^*\|_{1,n^i}$ will be large on the elements close to the ith detector.

(v) $|g_j|$ and $|g_i^*|$ are higher close to the sources and detectors, respectively. Therefore, h_m^j has to be small around the jth source and around all detectors, where α^λ is nonzero. Likewise, h_n^i has to be small around the ith detector and around all sources, where α^λ is nonzero.

(vi) If α^λ is nonzero on the whole domain Ω, then the error may become higher depending on the magnitude of $|g_j|$ and $|g_i^*|$.

(vii) The regularization parameter affects the bound on $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)}$.

(viii) Increasing the number of sources and detectors increases the bound on $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)}$.

Remark 4.

(i) Note that the finite-element discretization error estimates (3.8)–(3.9) depend on only the smoothness and support of g_j and g_i^*, and the finite dimensional space of approximating functions [9]. However, the error estimate (4.28) in theorem 2 shows that the accuracy of the reconstructed image $\tilde{\alpha}^\lambda$ depends on the orientation of the absorptive heterogeneity with respect to the sources and detectors, as well as on the finite-element discretization error estimates (3.8)–(3.9). In this respect, the estimate (4.28) in theorem 2 shows that reducing the discretization error in the solutions G_j and G_i^* of the forward problem may not ensure the accuracy of the reconstructed absorption image (see [14]).

(ii) In case a different discretization approach such as finite difference [20] or finite volume [13] is used to solve the forward problem, theorem 2 can be modified in a straightforward manner by replacing the discretization error estimates (3.8) and (3.9) with the corresponding error estimates specific to the method of choice [13, 20].

(iii) Let $\tilde{\gamma}^\delta$ be the perturbed right-hand side $\tilde{\gamma}$ of (4.26) due to the presence of noise, such that $\|\tilde{\gamma}^\delta - \tilde{\gamma}\|_{L^1(\Omega)} \leqslant \tilde{\delta}$. Then, an additional term is introduced to the bound in (4.28) due to this perturbation:

$$\begin{aligned}\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)} \leqslant{}& \frac{C}{\lambda}\max_{i,j}\|g_i^* g_j\|_{L^1(\Omega)}\\ &\times\left(\sum_{i=1}^{N_d}\sum_{n,j}^{N_\Delta^{*i},N_s}(2\|g_j\alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty\|g_j\|_{0,n^i})\|g_i^*\|_{1,n^i}h_n^i\right.\\ &\left.+\sum_{j=1}^{N_s}\sum_{m,i}^{N_\Delta^j,N_d}(2\|g_i^*\alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty\|g_i^*\|_{0,m^j})\|g_j\|_{1,m^j}h_m^j\right) + \frac{\tilde{\delta}}{\lambda}.\end{aligned} \tag{4.35}$$

Clearly, the additional term $\tilde{\delta}/\lambda$ due to the presence of noise in (4.35) is independent of the discretization of the forward problem.

(iv) Theorem 2 provides a general framework to analyse the error in reconstructed optical images resulting from the perturbations in the kernel of the linear integral equation (2.16). In general, a perturbation in the kernel of the linear integral equation (2.16) can occur due to errors resulting from the numerical integration of (2.6)–(2.7), the approximation of the boundary $\partial\Omega$, the inaccurate approximation of the source Q_j and/or the background optical properties. Furthermore, the analysis framework in theorem 2 can be used to analyse the effect of linearization of the Lippmann–Schwinger-type equations [10] using Born approximation on the accuracy of the reconstructed optical images [15].

(v) A bound similar to (4.28) can be derived for the error that occurs as a result of the discretization of the forward problem in electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography and microwave imaging.

4.3. Iterative Born approximation

In this section, we explore the error in the inverse problem solution within an iterative linearization approach.

The error analysis presented in this paper covers the error which results from the discretization of the forward and inverse problems. If α is sufficiently low, then one iteration suffices to solve the inverse problem and the error analysis discussed above applies. When iterative linearization is considered to address the nonlinearity of the inverse problem, we can make use of the error analysis at each linearized step as follows: let $\alpha^{\lambda}_{(t)}$ and $\tilde{\alpha}^{\lambda}_{n(t)}$ be the actual solution of the regularized inverse problem (2.13) and the solution of (3.10) at the tth linearization step, respectively. At the end of the $(r-1)$th linearization step, the absorption coefficient estimate at $\mathbf{x}$ is given by $\hat{\mu}_a^{(r-1)}(\mathbf{x}) = \mu_a^{(0)}(\mathbf{x}) + \sum_{t=1}^{r-1} \tilde{\alpha}^{\lambda}_{n(t)}(\mathbf{x})$, where $\tilde{\alpha}^{\lambda}_{n(t)}$ has an error due to discretization with respect to the actual solution $\alpha^{\lambda}_{(t)}$, and $\mu_a^{(0)}$ is the initial guess for the background absorption coefficient. In the next linearization, an error on the new solution update $\hat{\mu}_a^{(r)}$ will be introduced due to

(i) projection (inverse problem discretization),

(ii) the error $(\tilde{\mathcal{K}} - \mathcal{K})^{(r-1)}$ in the operator $(\tilde{\mathcal{K}})^{(r-1)}$ and the error $(\tilde{\gamma} - \gamma)^{(r-1)}$ in $(\tilde{\gamma})^{(r-1)}$ resulting from the forward problem discretization, and

(iii) the error in the $(r-1)$th update $\hat{\mu}_a^{(r-1)}$, resulting from the discretization of the forward and inverse problems. Note that $\hat{\mu}_a^{(r-1)}$ appears as a coefficient in the boundary value problems (2.1)–(2.2) and (2.3)–(2.4). An error in this coefficient implies perturbation in the solutions of (2.1)–(2.2) and (2.3)–(2.4). As a result, G_j and G_i^* will have error terms in addition to the discretization error.

As a result, the error in $\hat{\mu}_a^{(r)}$ at the rth iteration is bounded by

$$\left\|\mu_a - \hat{\mu}_a^{(r)}\right\| = \left\|\sum_{t=1}^{r} \alpha^{\lambda}_{(t)} - \tilde{\alpha}^{\lambda}_{n(t)}\right\| \leqslant \sum_{t=1}^{r} \left\|\alpha^{\lambda}_{(t)} - \tilde{\alpha}^{\lambda}_{n(t)}\right\|, \tag{4.36}$$

assuming that the initial guess $\mu_a^{(0)}$ for the background absorption is approximated accurately while solving the boundary value problems (2.1)–(2.2) and (2.3)–(2.4) at the first iteration, that is $\mu_a^{(0)}(\mathbf{x}) - \sum_{k=1}^{n} \mu_a^{(0)}(\mathbf{x}_k) L_k(\mathbf{x}) \to 0$, for all $\mathbf{x} \in \Omega$.

5. Conclusion

In this work, we presented an error analysis to show the relationship between the error in the reconstructed optical absorption images and the discretization of the forward and inverse problems. We summarized the implications of the error analysis in two theorems which provide an insight into the impact of forward and inverse problem discretizations on the accuracy of the reconstructed optical absorption images. These theorems show that the error in the reconstructed optical image due to the discretization of each problem is bounded by roughly the multiplication of the discretization error in the corresponding solution and the solution of the other problem. In particular, theorem 2 shows that solving the diffusion equation and the associated adjoint problem accurately may not ensure small values for $\|\tilde{\mathcal{K}} - \mathcal{K}\|$ and $\|\gamma - \tilde{\gamma}\|$, which may lead to large errors in the reconstructed optical images, depending on the value of the regularization parameter. Similarly, relatively large discretization error in the solution of the forward problem may have relatively low impact on the accuracy of the reconstructed optical images, depending on the source–detector configuration, and orientation with respect to the optical heterogeneities. We have also shown that the error estimates can be extended to include the effect of noise on the overall error in the reconstructed images.

The error analysis presented in this work motivates the development of novel adaptive discretization schemes based on the error estimates in theorems 1 and 2. In the sequel of this work, we propose two novel adaptive discretization algorithms for the forward and inverse problems [14], and justify the validity of theorems 1 and 2.

The error analysis can be extended to show the effect of the discretization error on the accuracy of the simultaneous reconstruction of scattering and absorption coefficients, which will be the focus of our future work. Finally, we note that the error analysis introduced in this paper is not limited to DOT, and can easily be adapted for similar inverse parameter estimation problems such as electrical impedance tomography, bioluminescence tomography, optical fluorescence tomography, microwave imaging etc.

Acknowledgments

This material is based upon work supported by NSF-BES-0353160, ONR-N00014-04-1-0694, and US Army Medical Research-W81XWH-04-1-0559. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Various portions of this research were supported by the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (Award Number EEC-9986821) and Rensselaer Polytechnic Institute.

Appendix A. Boundedness of $\mathcal{A}_a$

$$\|\mathcal{A}_a\alpha\|_{l^1} = \sum_{i,j}^{N_d,N_s} \left| \int_\Omega H_{i,j}(\mathbf{x})\alpha(\mathbf{x})\,\mathrm{d}\mathbf{x} \right|. \tag{A.1}$$

We can write the following inequality:

$$\|\mathcal{A}_a\alpha\|_{l^1} \leqslant \sum_{i,j}^{N_d,N_s} \int_\Omega |H_{i,j}(\mathbf{x})\alpha(\mathbf{x})|\,\mathrm{d}\mathbf{x} \leqslant \left(\sum_{i,j}^{N_d,N_s} \int_\Omega |H_{i,j}(\mathbf{x})|\,\mathrm{d}\mathbf{x} \right) \|\alpha\|_\infty. \tag{A.2}$$

Using Schwarz' inequality, we can write an upper bound for the summation as follows:

$$\begin{aligned}\sum_{i,j}^{N_d,N_s}\int_\Omega |H_{i,j}(\mathbf{x})|\,\mathrm{d}\mathbf{x} &= \sum_{i,j}^{N_d,N_s}\|g_i^* g_j\|_{L^1(\Omega)}\\ &\leqslant \sum_{i,j}^{N_d,N_s}\|g_i^*\|_0\|g_j\|_0\\ &\leqslant N_d N_s \max_i\|g_i^*\|_0 \max_j\|g_j\|_0,\end{aligned} \tag{A.3}$$

which leads to

$$\|\mathcal{A}_a\alpha\|_{l^1} \leqslant N_d N_s \max_i\|g_i^*\|_0 \max_j\|g_j\|_0\|\alpha\|_\infty.$$

Therefore an upper bound for the norm of $\mathcal{A}_a$ is given by

$$\|\mathcal{A}_a\|_{L^\infty(\Omega)\to l^1} \leqslant N_d N_s \max_i\|g_i^*\|_0 \max_j\|g_j\|_0. \tag{A.4}$$

The boundedness of g_j and g_i^* imply that $\mathcal{A}_a$ is bounded.

Appendix B. Compactness of $\mathcal{A}_a$

$\mathcal{A}_a$ is bounded by (A.4). Furthermore $\mathcal{A}_a$ maps the infinite-dimensional subspace $L^\infty(\Omega)$ to a finite-dimensional subspace $\mathbb{C}^{N_d\times N_s}$, that is the range $R(\mathcal{A}_a)$ of $\mathcal{A}_a$ satisfies $R(\mathcal{A}_a) \in \mathbb{C}^{N_d\times N_s}$ due to the finite number of sources and detectors. As a result, $\mathcal{A}_a$ is compact [18]. The inverse problem is illposed as a consequence of compactness [18].

Appendix C. Proof of the lemma

The identity operator $\mathcal{I}$ is a bounded operator with bounded inverse and $(\mathcal{P}_n\mathcal{I})^{-1} = \mathcal{I} : X_n \to X_n$. Furthermore, $\|\mathcal{P}_n\|_{X\to X_n}$ is bounded for first-order Lagrange basis functions [4, 18]. Thus, projection by collocation converges for the identity operator. $\mathcal{A}$ is bounded and compact, and $\mathcal{K} = \lambda\mathcal{I} + \mathcal{A}$ is injective, with bounded inverse given by (4.1). As a result, by theorem 13.7 in [18], the projection method also converges for $\mathcal{K} = \lambda\mathcal{I} + \mathcal{A}$. Convergence of projection for $\mathcal{K}$ implies $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\alpha^\lambda \to \alpha^\lambda, n \to \infty$ for $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K} : X \to X_n$ [18].

It follows from the proof of theorem 13.7 in [18] that $\left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)^{-1} : Y_n \to X_n$ exists and is uniformly bounded for all sufficiently large n. Then from $\mathcal{P}_n\mathcal{K} = \lambda\mathcal{P}_n\left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right) = \lambda\left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)$, it follows that $\mathcal{P}_n\mathcal{K} : X_n \to Y_n$ is invertible for all sufficiently large n with the inverse given by

$$(\mathcal{P}_n\mathcal{K})^{-1} = \left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)^{-1}\frac{1}{\lambda}. \tag{C.1}$$

As a result we can write $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}$ as follows:

$$(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K} = \left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)^{-1}\frac{1}{\lambda}\mathcal{P}_n\mathcal{K}. \tag{C.2}$$

Thus,

$$\|(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\|_{X\to X_n} \leqslant C_M\frac{\|\mathcal{K}\|_{X\to Y}}{\lambda} \tag{C.3}$$

where $C_M > 0$ is independent of n, using the facts that projection by collocation method converges for the identity operator and $\left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}\right)^{-1}$ is uniformly bounded.

References

[1] Ainsworth M and Oden J T 1993 A unified approach to *a posteriori* error estimation using elemental residual methods *Numer. Math.* **65** 23–50

[2] Arridge S R 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41–93

[3] Arridge S R, Kaipio J P, Kolehmainen V, Schweiger M, Somersalo E, Tarvainen T and Vauhkonen M 2006 Approximation errors and model reduction with an application in optical diffusion tomography *Inverse Problems* **22** 175–95

[4] Atkinson K and Weimin H 2001 *Theoretical Numerical Analysis: A Functional Analysis Framework* (Berlin: Springer)

[5] Babuška I and Rheinboldt W C 1978 Error estimates for adaptive finite element computations *SIAM J. Numer. Anal.* **15** 736–54

[6] Babuška I, Zienkiewicz O C, Gago J and Oliveira E R de A 1986 *Accuracy Estimates and Adaptive Refinements in Finite Element Computations* (New York: Wiley)

[7] Bank R E and Weiser A 1985 Some *a posterior* error estimators for elliptic partial differential equations *Math. Comput.* **44** 283–301

[8] Beilina L 2002 Adaptive hybrid FEM/FDM methods for inverse scattering problems *PhD Thesis* Chalmers University of Technology

[9] Brenner S C and Scott L R 2002 *The Mathematical Theory of Finite Element Methods* (Berlin: Springer)

[10] Colton D and Kress R 1998 *Inverse Acoustic and Electromagnetic Scattering Theory* (Berlin: Springer)

[11] Daners D 2000 Robin boundary value problems on arbitrary domains *Trans. Am. Math. Soc.* **352** 4207–36

[12] Evans L C 1998 *Partial Differential Equations* (Providence, RI: American Mathematical Society)

[13] Eymard R, Gallouët T and Herbin R 2000 The finite volume method *Handbook of Numerical Analysis* vol 7 ed P G Ciarlet and J L Lions (Amsterdam: Elsevier) pp 715–1022

[14] Guven M, Yazici B, Kwon K, Giladi E and Intes X 2007 Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: II *Inverse Problems* **23** 1135–60

[15] Guven M, Yazici B, Kwon K, Giladi E and Intes X Error in optical absorption images due to Born approximation in diffuse optical tomography, in preparation

[16] Ihlenburg F 1998 *Finite Element Analysis of Acoustic Scattering* (Berlin: Springer)

[17] Kaipio J and Somersalo E 2005 Statistical and computational inverse problems *Applied Mathematical Sciences* vol 160 (New York: Springer)

[18] Kress R 1999 Linear integral equations *Applied Mathematical Sciences* vol 82 2nd edn (Berlin: Springer)

[19] Li R, Liu W, Ma H and Tang T 2002 Adaptive finite element approximation for distributed elliptic optimal control problems *SIAM J. Control Optim.* **41** 1321–49

[20] Smith G D 1985 *Numerical Solution of Partial Differential Equations: Finite Difference Methods (Oxford Applied Math. Comput. Sci. Series)* 3rd edn (Oxford: Clarendon)

[21] Strouboulis T and Hague K A 1992 Recent experiences with error estimation and adaptivity: Part I. Review of error estimators for scalar elliptic problems *Comput. Methods Appl. Mech. Eng.* **97** 399–436

[22] Verfurth R 1996 *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques* (New York: Teubner-Wiley)

IOP PUBLISHING INVERSE PROBLEMS

Inverse Problems **23** (2007) 1135–1160 doi:10.1088/0266-5611/23/3/018

Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: II

Murat Guven[1], **Birsen Yazici**[1], **Kiwoon Kwon**[1], **Eldar Giladi**[2] **and Xavier Intes**[3]

[1] Electrical, Computer, and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY, USA
[2] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, USA
[3] Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA

E-mail: yazici@ecse.rpi.edu

Received 28 September 2006
Published 15 May 2007
Online at stacks.iop.org/IP/23/1135

Abstract

In part I (Guven *et al* 2007 *Inverse Problems* **23** 1115–33), we analysed the error in the reconstructed optical absorption images resulting from the discretization of the forward and inverse problems. Our analysis led to two new error estimates, which present the relationship between the optical absorption imaging accuracy and the discretization error in the solutions of the forward and inverse problems. In this work, based on the analysis presented in part I, we develop new adaptive discretization schemes for the forward and inverse problems in order to reduce the error in the reconstructed images resulting from discretization. The proposed discretization schemes lead to adaptively refined composite meshes that yield the desired level of imaging accuracy while reducing the size of the discretized forward and inverse problems. We present numerical experiments to validate the error estimates developed in part I and show the improvement in the accuracy of the reconstructed optical images with the new adaptive mesh generation algorithms.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Numerical approaches in solving the forward and inverse problems in diffuse optical tomography (DOT) pose a tradeoff between computational efficiency and imaging accuracy. This tradeoff is a direct consequence of the discretization of the forward and inverse problems [2, 9] and the size of the resulting discrete forward and inverse problems. The imaging accuracy depends on the discretization error in the forward and inverse problem solutions. On the other hand, attempting to minimize the discretization error in the solutions of both problems separately implies a significant increase in the size of the discrete forward and inverse problems. Hence, it is important to understand the relationship between the discretization error and the resulting error in the solution of the inverse problem. Such a relationship can illuminate the mutual dependence of the forward and inverse problem solutions and identify the factors that control the extent to which the discretization error in the solutions of the forward and inverse problems affects the accuracy of the reconstructed optical images.

In part I of this two-part study, we presented an error analysis which showed the effect of discretization of the forward and inverse problems on the accuracy of the reconstructed optical absorption images [9]. The analysis led to two new error estimates that took into account the interdependence of the forward and inverse problems (see section 2). In the second part of our work, based on the error analysis presented in part I, we develop new adaptive discretization schemes for the forward and inverse problems. The resulting locally refined meshes reduce the error in the reconstructed optical images while keeping the size of the discrete forward and inverse problems relatively small.

There has been extensive research on adaptive mesh generation for the numerical solution of partial differential equations (see [9] for a list of publications) and inverse parameter estimation problems to reduce the undesired effect of discretization error [4, 14]. In the area of DOT, in [3] it was numerically shown that approximation errors resulting from the discretization of the forward problem can lead to significant degradation in the quality of the reconstructed images. In that work, the error in the reconstructed images is minimized by using an enhanced imaging model that treats this additional approximation error within the Bayesian framework. Alternatively, several investigators have reported on adaptive discretization schemes for the forward and inverse problems to address the optical image degradation due to discretization. In [6] a 'data-driven zonation' scheme, which can be viewed as an adaptive discretization algorithm, was proposed for fluorescence imaging [6]. In [8], we presented a region-of-interest (ROI) imaging scheme for DOT, which employed a multi-level algorithm on a non-uniform grid. The non-uniform grid is designed so as to provide finer spatial resolution for the ROI which corresponds to the tumour region as indicated by *a priori* anatomical image. In [16] an *a priori* non-uniform mesh design which provides high resolution at the heterogeneities and near boundary regions was proposed. In that work, the mesh refinement is independent of the source–detector configuration and the location of the heterogeneities. In [7] a dual mesh strategy was proposed, in which, a relatively fine uniform mesh is considered for the forward problem discretization and a coarse uniform mesh is generated for the inverse problem discretization. In the same study, an adaptive refinement scheme was proposed for the inverse problem discretization, but no adaptive refinement was considered for the solution of the forward problem. Another dual mesh strategy which makes use of *a priori* ultrasound information was presented in [10]. In that work, the dual mesh is a coarse mesh for the background tissue and a relatively fine mesh for the heterogeneity, similar to the approach in [8]. In fluorescence imaging, a dual adaptive mesh strategy was used to discretize the inverse problem and the associated coupled diffusion equations, where the refinement criterion is based on *a posteriori* discretization error estimates [12]. Note that in all these studies [6–8, 10, 12, 16], the mesh refinement criteria considered for the inverse (forward) problem disregard the impact of the solution of the forward (inverse) problem. In other words, the discretization of each problem is considered independently of the solution of the other problem.

In this work, based on the two error bounds provided by the error analysis in part I [9], we introduce an adaptive discretization scheme for the forward and inverse problems, respectively. We remark that the mesh refinement criterion for each problem comprises the discretization error in the corresponding problem solution, scaled spatially by the solutions of both problems.

Table 1. Definition of function spaces and norms.

Notation	Explanation
$\overline{f}$	The complex conjugate of the function f
$C(\Omega)$	Space of continuous complex-valued functions on Ω
$C^k(\Omega)$	Space of complex-valued k-times continuously differentiable functions on Ω
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \vert \operatorname{ess\,sup}_\Omega \lvert f(\mathbf{x})\rvert < \infty\}$
$L^p(\Omega)$	$L^p(\Omega) = \left\{ f \,\middle\vert\, \left(\int_\Omega \lvert f(\mathbf{x})\rvert^p \,\mathrm{d}\mathbf{x}\right)^{1/p} < \infty \right\}, p \in [1, \infty(b))$
$D_w^z f$	zth weak derivative of f
$H^p(\Omega)$	$H^p(\Omega) = \left\{ f \,\middle\vert\, \left(\sum_{\lvert z\rvert \leqslant p} \left\Vert D_w^z f \right\Vert_0^2\right)^{1/2} < \infty \right\}, p \in [1, \infty)$
$\Vert f\Vert_0$	The $L^2(\Omega)$ norm of f
$\Vert f\Vert_p$	The $H^p(\Omega)$ norm of f
$\Vert f\Vert_\infty$	The $L^\infty(\Omega)$ norm of f
$\Vert f\Vert_{L^p(\Omega)}$	The $L^p(\Omega)$ norm of f
$\Vert f\Vert_{0,m}$	The L^2 norm of f over the mth finite element Ω_m
$\Vert f\Vert_{p,m}$	The H^p norm of f over the mth finite element Ω_m

Thus, the proposed adaptive mesh generation algorithms address the interdependence between the solutions of the forward and inverse problems and take into account the orientation of the source–detectors and the absorptive perturbations. This makes the adaptive discretization algorithms introduced in this paper different from the previous approaches [6–8, 10, 12, 16]. The simulation experiments validate the implications of our error analysis and show that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems. We specifically show that using the discretization error estimates, which do not take into account the interdependence of forward and inverse problems as a criterion for discretization, may lead to severely degraded image reconstructions (see simulation study 3). We also discuss the computational complexity of the proposed adaptive mesh generation algorithms and compare it to the computational complexity of mesh generation algorithms based on the conventional discretization error estimates. We finally note that the proposed adaptive mesh generation algorithms can be adapted for similar inverse parameter estimation problems, such as electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, microwave imaging, etc.

The outline of this paper is as follows: in section 2, we give a brief overview of the forward and inverse DOT problems and recall the two theorems presented in part I which summarize the impact of discretization on the accuracy of the reconstructed optical images. In section 3, based on these two theorems, we introduce the adaptive mesh generation algorithms for the solution of the forward and inverse problems and discuss their computational complexity. In section 4, we present our experimental results, which is followed by section 5. The appendix includes the solution of a model problem used to initiate the adaptive mesh generation.

2. Overview

In this section, we first briefly define the forward and inverse problems in DOT. Next, we state theorems 1 and 2 presented in the first part of this work [9] to recall the effect of the discretization of the forward and inverse problems on the accuracy of optical absorption image reconstruction. We refer to table 1 for the explanation of the notation associated with functions and their norms. Note that calligraphic letters are used to denote the operators, e.g. $\mathcal{A}_a$, $\mathcal{I}$, $\mathcal{K}$ etc.

2.1. Forward and inverse problems in DOT

We consider the following boundary value problem to model the near-infrared light propagation in a bounded domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$ [2, 5]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_j(\mathbf{x}) + \left(\mu_a(\mathbf{x}) + \frac{\mathrm{i}\omega}{c}\right) g_j(\mathbf{x}) = Q_j(\mathbf{x}) \qquad \mathbf{x} \in \Omega, \tag{2.1}$$

$$g_j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_j}{\partial n}(\mathbf{x}) = 0 \qquad \mathbf{x} \in \partial\Omega, \tag{2.2}$$

where $g_j(\mathbf{x})$ is the photon density at $\mathbf{x}$, Q_j is the point source located at the source position $\mathbf{x}_s^j$, $j = 1, \ldots, N_s$, where N_s is the number of sources, $D(\mathbf{x})$ is the diffusion coefficient and $\mu_a(\mathbf{x})$ is the absorption coefficient at $\mathbf{x}$, $\mathrm{i} = \sqrt{-1}$, ω is the modulation frequency of the source, c is the speed of the light, $a = (1 + R)/(1 - R)$ where R is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial \cdot /\partial n$ denotes the directional derivative along the unit normal vector on the boundary. The boundary value problem (2.1)–(2.2) constitutes the forward problem in DOT together with the associated adjoint problem [2, 9]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^*(\mathbf{x}) + \left(\mu_a(\mathbf{x}) - \frac{\mathrm{i}\omega}{c}\right) g_i^*(\mathbf{x}) = 0 \qquad \mathbf{x} \in \Omega, \tag{2.3}$$

$$g_i^*(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_i^*}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \qquad \mathbf{x} \in \partial\Omega, \tag{2.4}$$

where Q_i^* is the adjoint source located at the detector position $\mathbf{x}_d^i$, $i = 1, \ldots, N_d$, where N_d is the number of detectors. Note that we approximate the point source Q_j in (2.1) and the adjoint source Q_i^* in (2.4) by Gaussian functions with sufficiently low variance, whose centres are located at $\mathbf{x}_s^j$ and $\mathbf{x}_d^i$, respectively.

In this work, we focus on the estimation of the absorption coefficient and consider an iterative algorithm based on repetitive linearization of the inverse problem using first-order Born approximation. Using a zeroth-order Tikhonov regularization to address the illposedness, the inverse problem at each iteration reads

$$\begin{aligned}
\gamma(\mathbf{x}) &:= (\mathcal{A}_a^*\Gamma)(\mathbf{x}) = [(\mathcal{A}_a^*\mathcal{A}_a + \lambda\mathcal{I})\alpha^\lambda](\mathbf{x}) \\
&:= \int_\Omega \kappa(\mathbf{x}, \acute{\mathbf{x}})\alpha^\lambda(\acute{\mathbf{x}})\,\mathrm{d}\acute{\mathbf{x}} + \lambda\alpha^\lambda(\mathbf{x}) \qquad (2.5)\\
&:= (\mathcal{K}\alpha^\lambda)(\mathbf{x}), \qquad (2.6)
\end{aligned}$$

where $\Gamma \in \mathbb{C}^{N_d \times N_s}$ is the vector of differential measurements at N_d number of detectors due to N_s number of sources, as a result of the small perturbation α on the background absorption coefficient μ_a, and α^λ is the solution of the regularized inverse problem. In (2.5), $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ is the kernel of the integral equation, given by [9]

$$\kappa(\mathbf{x}, \acute{\mathbf{x}}) = \sum_{i,j}^{N_d, N_s} H_{i,j}^*(\mathbf{x}) H_{i,j}(\acute{\mathbf{x}}), \tag{2.7}$$

where $H_{i,j} := -\overline{g_i^*}g_j$ is the (i, j)th kernel of the matrix-valued operator $\mathcal{A}_a : L^\infty(\Omega) \to \mathbb{C}^{N_d\times N_s}$ and $H_{i,j}^* := -g_i^*\overline{g_j}$ is the (i, j)th kernel of the adjoint operator $\mathcal{A}_a^* : \mathbb{C}^{N_d\times N_s} \to L^1(\Omega)$ defined by

$$(\mathcal{A}_a^*\beta)(\mathbf{x}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x})\beta_{i,j} = \sum_{i,j}^{N_d,N_s} -g_i^*(\mathbf{x})\overline{g_j(\mathbf{x})}\beta_{i,j}, \tag{2.8}$$

for all $\beta \in \mathbb{C}^{N_d\times N_s}$. We note that g_j and g_i^* in (2.7) and (2.8) are the solutions of the variational formulations of (2.1)–(2.2) and (2.3)–(2.4), respectively [9]. Assume that $D, \mu_a \in C^1(\Omega)$. Noting $Q_j, Q_i^* \in H^1(\Omega)$, the solutions $g_j, g_i^* \in H^1(\Omega)$ of the variational formulations of the boundary value problems (2.1)–(2.2) and (2.3)–(2.4) also satisfy [9]

$$g_j, g_i^* \in C(\Omega). \tag{2.9}$$

For the rest of the paper, we will denote $L^\infty(\Omega)$ and $L^1(\Omega)$ by X and Y, respectively.

Below we summarize the two theorems of part I [9] and provide the error estimates which will be used in the design of adaptive meshes for the discretization of the forward and inverse DOT problems. In this respect, we first consider the impact of the inverse problem discretization when the associated kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ in (2.5) is exact. Next, we give the error estimate for the case in which the kernel is replaced by its finite-dimensional approximation (i.e. degenerate kernel) and analyse the effect of the forward problem discretization on the accuracy of the reconstructed image without projecting (2.6).

2.2. *Effect of inverse problem discretization*

Let $X_n \subset X$ and $Y_n \subset Y$ denote a sequence of finite dimensional subspaces of dimension $n = 1, 2, \ldots,$ spanned by first-order Lagrange basis functions $\{L_1, \ldots, L_n\}$, and $\{\mathbf{x}_p\}$, $p = 1, \ldots, n$, be the set of collocation points on Ω. Then, the discretization of the inverse problem (2.6) by projecting it onto the finite dimensional subspace Y_n using the collocation method approximates the solution of (2.6) by an element $\alpha_n^\lambda \in X_n$ which satisfies

$$\left(\mathcal{K}\alpha_n^\lambda\right)(\mathbf{x}_p) = \gamma(\mathbf{x}_p), \qquad p = 1, \ldots, n, \tag{2.10}$$

where we express $\alpha_n^\lambda(\mathbf{x})$, $\mathbf{x} \in \Omega$ on a set $\{\Omega_m\}$ of finite elements for $m = 1, \ldots, N_\Delta$ such that $\bigcup_m^{N_\Delta} \Omega_m = \Omega$ as follows:

$$\alpha_n^\lambda(\mathbf{x}) = \sum_{k=1}^{n} a_k L_k(\mathbf{x}). \tag{2.11}$$

Equivalently, the collocation method can be interpreted as a projection with the interpolation operator $\mathcal{P}_n : Y \to Y_n$ defined by [13]

$$\mathcal{P}_n f(\mathbf{x}) := \sum_{p=1}^{n} f(\mathbf{x}_p)L_p(\mathbf{x}), \qquad \mathbf{x} \in \Omega, \tag{2.12}$$

for all $f \in Y$. Then, (2.10) is equivalent to

$$\mathcal{P}_n\mathcal{K}\alpha_n^\lambda = \mathcal{P}_n\gamma. \tag{2.13}$$

Let ψ be the interpolant of α^λ [5] and assume that $\alpha^\lambda \in H^1(\Omega)$. Then, the interpolation error $e_\alpha = \alpha^\lambda - \psi$ on each finite element Ω_m is bounded by

$$\|e_\alpha\|_{0,m} \leqslant C\|\alpha^\lambda\|_{1,m}h_m, \tag{2.14}$$

where C is a positive constant and h_m is the diameter of the smallest ball that contains the mth element Ω_m.

Theorem 1 describes the effect of inverse problem discretization on the accuracy of the reconstructed optical absorption image.

Theorem 1. *Let g_j, g_i^* be the solutions of the variational formulations of the boundary value problems (2.1)–(2.2) and (2.3)–(2.4), respectively. The error between the solution α^λ of (2.6) and the solution α_n^λ of (2.13) is bounded by*

$$\begin{aligned}\left\|\alpha^\lambda - \alpha_n^\lambda\right\|_{L^1(\Omega)} &\leqslant C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\|\alpha^\lambda\|_{1,m}h_m \\ &\quad + \frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\max_{i,j}\|g_i^* g_j\|_{L^1(\Omega)}\sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s}\|g_i^* g_j\|_{0,m}\|\alpha^\lambda\|_{1,m}h_m,\end{aligned}\tag{2.15}$$

where C is a positive constant, V_Ω is the volume of Ω, $\mathcal{T}_n : Y \to X_n$ is a uniformly bounded operator given by $\mathcal{T}_n = \left(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}_a^*\mathcal{A}_a\right)^{-1}\mathcal{P}_n$ [9].

Proof. See [9]. □

2.3. Effect of forward problem discretization

Let $\left\{\Omega_m^j\right\}$ denote the set of linear elements used to discretize the variational formulation of the boundary value problem (2.1)–(2.2) for $m = 1, \ldots, N_\Delta^j$; such that $\bigcup_m^{N_\Delta^j}\Omega_m^j = \Omega$, and h_m^j be the diameter of the smallest ball that contains the element Ω_m^j in the finite-dimensional solution G_j, for all $j = 1, \ldots, N_s$ [9]. Similarly, let $\left\{\Omega_n^i\right\}$ denote the set of linear elements used to discretize the variational formulation of the boundary value problem (2.3)–(2.4) for $n = 1, \ldots, N_\Delta^{*i}$; such that $\bigcup_n^{N_\Delta^{*i}}\Omega_n^i = \Omega$, and h_n^i be the diameter of the smallest ball that contains the element Ω_n^i in the finite-dimensional solution G_i^*, for all $i = 1, \ldots, N_d$ [9]. Then, a bound for the discretization error in the finite element solutions G_j and G_i^* with respect to the solutions g_j and g_i^* of the variational formulations of the boundary value problems (2.1)–(2.2) and (2.3)–(2.4) on each finite element can be given by [5]

$$\|g_j - G_j\|_{0,m^j} \leqslant C\|g_j\|_{1,m^j}h_m^j,\tag{2.16}$$

$$\|g_i^* - G_i^*\|_{0,n^i} \leqslant C\|g_i^*\|_{1,n^i}h_n^i,\tag{2.17}$$

where C is a positive constant, and $\|\cdot\|_{0,m^j}$ ($\|\cdot\|_{0,n^i}$) and $\|\cdot\|_{1,m^j}$ ($\|\cdot\|_{1,n^i}$) are respectively the L^2 and H^1 norms on $\Omega_m^j\left(\Omega_n^i\right)$.

Consider the inverse problem

$$\tilde{\mathcal{K}}\tilde{\alpha}^\lambda = \tilde{\gamma},\tag{2.18}$$

where $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ are the finite dimensional approximations to $\mathcal{K}$ and γ, obtained by substituting g_j and g_i^* in $H_{i,j}$ and $H_{i,j}^*$ by G_j and G_i^*, respectively.

Theorem 2 shows the effect of forward problem discretization on the accuracy of the reconstructed optical absorption image.

Theorem 2. *A bound for the error between the solution α^λ of (2.6) and the solution $\tilde{\alpha}^\lambda$ of (2.18) due to approximations $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ is given by*

$$\|\alpha^{\lambda} - \tilde{\alpha}^{\lambda}\|_{L^1(\Omega)} \leqslant \frac{C}{\lambda} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} (2\|g_j \alpha^{\lambda}\|_{0,n^i} + \|\alpha\|_\infty \|g_j\|_{0,n^i}) \|g_i^*\|_{1,n^i} h_n^i \right.$$
$$\left. + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^{j}, N_d} (2\|g_i^* \alpha^{\lambda}\|_{0,m^j} + \|\alpha\|_\infty \|g_i^*\|_{0,m^j}) \|g_j\|_{1,m^j} h_m^j \right), \tag{2.19}$$

where C is a positive constant.

Proof. See [9]. ☐

We refer to part I [9] for further details of the discussion regarding the definition and discretization of the forward and inverse problems. In the following, we discuss the adaptive mesh generation for the forward and inverse problems.

3. Adaptive mesh generation

In this section, we discuss the adaptive mesh design for the discretization of the forward and inverse problems based on theorems 1 and 2. For each problem, we present an adaptive mesh generation algorithm, which is followed by the corresponding computational cost analysis.

3.1. Adaptive mesh generation for the forward problem

Let the mesh parameter h_m^j for G_j, $j = 1, \ldots, N_s$, and the mesh parameter h_n^i for G_i^*, $i = 1, \ldots, N_d$ be chosen so that

$$h_m^j \leqslant \frac{\epsilon_f}{\sum_{i=1}^{N_d} (2\|g_i^* \alpha^{\lambda}\|_{0,m^j} + \|\alpha\|_\infty \|g_i^*\|_{0,m^j}) \|g_j\|_{1,m^j}} := B_j^m, \tag{3.20}$$

$$h_n^i \leqslant \frac{\epsilon_f}{\sum_{j=1}^{N_s} (2\|g_j \alpha^{\lambda}\|_{0,n^i} + \|\alpha\|_\infty \|g_j\|_{0,n^i}) \|g_i^*\|_{1,n^i}} := B_i^{*n}, \tag{3.21}$$

where the tolerance ϵ_f will be defined later. Then, by theorem 2, the error in the reconstructed image due to the forward problem discretization is bounded by

$$\frac{C}{\lambda} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \left(\sum_{j=1}^{N_s} N_\Delta^j + \sum_{i=1}^{N_d} N_\Delta^{*i} \right) \epsilon_f = \tilde{\epsilon}^f, \tag{3.22}$$

where C is a positive constant and $\tilde{\epsilon}^f$ is the total allowable error in the reconstructed optical image due to the forward problem discretization. Equation (3.22) implies the following value for ϵ_f:

$$\epsilon_f = \frac{\lambda \tilde{\epsilon}^f / C}{\max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)}} \frac{1}{\left(\sum_{j=1}^{N_s} N_\Delta^j + \sum_{i=1}^{N_d} N_\Delta^{*i} \right)}. \tag{3.23}$$

Algorithm 1 outlines the adaptive mesh generation algorithm for the forward problem in the form of a pseudocode. The algorithm is performed for each source and detector before the linearization of the inverse problem and it yields a family of adaptively refined meshes with conforming elements. We use Rivara's algorithm [15] for refinement.

Algorithm 1. The pseudocode for the mesh generation algorithm for the forward problem, prior to the linearization of the inverse problem.

```
◇ Generate an initial uniform mesh $(\Delta, N_\Delta)$, $\Delta = \bigcup_{e=1}^{N_\Delta}\{\Delta_e\}$
◇ Set $\epsilon^f$
◇ Initialize the set of marked elements: $M_e \leftarrow \{\}$
◇ flag = True
  while flag = True
    for each element $\Delta_e \in \Delta$ with mesh parameter $h_e^j$ $\left(h_e^i\right)$
      if first linearization
        . Use analytical solutions for $g_j$ and $g_i^*$ and a priori anatomical
        information about $\alpha$ to compute the bound $B_j^m$ in (3.20) ($B_i^{*n}$ in (3.21))
      else
        . Use current solution updates $G_j$ and $G_i^*$ and $\tilde{\alpha}_n^\lambda$
        to compute $B_j^m$ in (3.20) ($B_i^{*n}$ in (3.21))
      end
      if $h_e^j > B_j^m$ $\left(h_e^i > B_i^{*n}\right)$
        . $M_e \leftarrow M_e \bigcup \{\Delta_e\}$
      end
    end
    if $M_e \neq \{\}$
      . Refine the marked elements and update the mesh $\Delta$
      . $M_e \leftarrow \{\}$
    else
      . flag = False
    end
  end
◇ Solve for $G_j$ $(G_i^*)$.
```

Remark 1.

(i) In practice, B_j^m and B_i^{*n} in (3.20)–(3.21) cannot be computed since α, α^λ, g_j and g_i^* are unknown. However, B_j^m and B_i^{*n} can be estimated by using approximations for the functions involved in these bounds, based on either *a priori* information or on the recent forward and inverse problem solution updates. Then, the elements whose mesh parameter h_m^j $\left(h_n^i\right)$ exceeds B_j^m $\left(B_i^{*n}\right)$ can be determined and refined.

(ii) After the first sweep of refinement, one can compute the bound B_j^m and B_i^{*n} only for the new elements. We note that for the initial mesh design, we use a model problem to compute the terms in the error bound relevant to the forward problem solution (see appendix). If there is no *a priori* information, α^λ can be assumed to be spatially constant at the first linearization step. After the first linearization, the norms in B_j^m and B_i^{*n} relevant to g_j and g_i^* are not expected to change significantly. In this context, the terms $\|g_i^*\alpha^\lambda\|_{0,m^j}$, $\|g_j\alpha^\lambda\|_{0,n^i}$ in (3.20) and (3.21) can be bounded by $\|g_i^*\|_{0,m^j}\|\alpha^\lambda\|_{\infty,m^j}$ and $\|g_j\|_{0,n^i}\|\alpha^\lambda\|_{\infty,n^i}$, respectively. Therefore, one can store the norms $\|g_j\|_{0,n^i}$ and $\|g_i^*\|_{0,m^j}$ at the end of the first mesh generation, and update B_j^m and B_i^{*n} in the following mesh generations by using these stored values and the updated α^λ values.

(iii) In case ϵ_f cannot be chosen in prior, we consider a posterior approach, set $\epsilon_f = 1$, and compute $h_m^j / B_j^m \left(h_n^i / B_i^{*n}\right)$ on each element, which is used as the indicator for refinement. Then, the elements with indicator value exceeding the average $h_m^j / B_j^m \left(h_n^i / B_i^{*n}\right)$ quantity are marked for refinement. We note that in this case, the algorithm has to be stopped when the number of nodes in the mesh exceeds the allowable number of nodes.

3.2. Computational cost of the adaptive mesh generation algorithm for the forward problem

Consider the algorithm described in remark 1(iii) for $\Omega \subset \mathbb{R}^2$. Using triangular finite elements with first-order Lagrange basis functions and an analytical (exact) integration on each finite element, the number of multiplications required to compute the L^2 or H^1 norm of a finite-dimensional function on each triangular element $\Omega_m^j \left(\Omega_n^i\right)$ is 12. On the other hand, computing the norm $\|g_j \alpha^\lambda\|_{0,n^i}$ $(\|g_i^* \alpha^\lambda\|_{0,m^j})$ takes ten times the number of multiplications to compute $\|g_j\|_{1,m^j}$ $(\|g_i^*\|_{1,n^i})$. As a result, the total number of multiplications required to compute the error estimates on all finite elements for the jth source is given by $(132N_d + 16)N_\Delta^j$. Similarly, the total number of multiplications required to compute the error estimates on all finite elements for the ith detector is equal to $(132N_s + 16)N_\Delta^{*i}$.

In order to reduce the computational cost of the proposed adaptive mesh generation algorithm, we can approximate the bounds B_j^m in (3.20) and B_i^{*n} in (3.21) as follows:

$$B_j^m \approx \frac{1}{\left(2\left\| \sum_{i=1}^{N_d} g_i^* \alpha^\lambda \right\|_{0,m^j} + \|\alpha\|_\infty \left\| \sum_{i=1}^{N_d} g_i^* \right\|_{0,m^j}\right) \|g_j\|_{1,m^j}} \tag{3.24}$$

$$B_i^{*n} \approx \frac{1}{\left(2\left\| \sum_{j=1}^{N_s} g_j \alpha^\lambda \right\|_{0,n^i} + \|\alpha\|_\infty \left\| \sum_{j=1}^{N_s} g_j \right\|_{0,n^i}\right) \|g_i^*\|_{1,n^i}}. \tag{3.25}$$

Then, the number of multiplications required to compute the error estimates on all finite elements becomes $148N_\Delta^j \left(148N_\Delta^{*i}\right)$, which implies a significant reduction as compared to $(132N_d + 16)N_\Delta^j \left((132N_s + 16)N_\Delta^{i*}\right)$.

If one uses the discretization error estimates (2.16)–(2.17) to generate adaptive meshes for the discretization of (2.1)–(2.2) and (2.3)–(2.4), the number of multiplications is equal to $13N_\Delta^j$ and $13N_\Delta^{*i}$, respectively. Then, the resulting adaptive meshes will lead to finite element solutions G_j and G_i^* with reduced discretization error. However, reduction in the discretization error in G_j and G_i^* may not ensure the accuracy of the reconstructed absorption image (see simulation experiment 3).

3.3. Adaptive mesh generation for the inverse problem:

Let the mesh parameter h_m for the solution of the inverse problem be defined as follows:

$$\begin{aligned} h_m \leqslant \epsilon_{\text{inv}} \Bigg/ \Bigg(&\sqrt{V_\Omega} \|\mathcal{I} - \mathcal{T}_n\|_{Y \to X_n} \|\alpha^\lambda\|_{1,m} + \frac{1}{\lambda} \|\mathcal{T}_n\|_{Y \to X_n} \\ &\times \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \sum_{i,j}^{N_d, N_s} \|g_i^* g_j\|_{0,m} \|\alpha^\lambda\|_{1,m} \Bigg) := B_{\text{inv}}^m. \end{aligned} \tag{3.26}$$

Then, by theorem 1, the error in the reconstructed image due to inverse problem discretization is bounded by

$$C N_\Delta \epsilon_{\text{inv}} = \tilde{\epsilon}^{\text{inv}}, \tag{3.27}$$

where C is a positive constant and $\tilde{\varepsilon}^{\text{inv}}$ is the total allowable error in the reconstructed optical image due to inverse problem discretization.

We present the pseudocode for our adaptive mesh generation algorithm used at each linearization of the inverse problem in algorithm 2. Similar to the forward problem discretization, we use Rivara's algorithm [15] for the refinement of the elements.

Algorithm 2. The pseudocode for the mesh generation algorithm at every linearization step of the inverse problem.

```
◇ Generate an initial uniform mesh $(\Delta, N_\Delta)$, $\Delta = \bigcup_{m=1}^{N_\Delta}\{\Delta_m\}$
◇ Set $\epsilon_{inv}$
◇ Initialize the set of marked elements: $M_e \leftarrow \{\}$
◇ flag = True
  while flag = True
    for each element $\Delta_m \in \Delta$ with mesh parameter $h_m$
      if first linearization
        . Use current solution updates $G_j$ and $G_i^*$ and a priori information
        about $\alpha$ to compute $B_{\text{inv}}^m$ in (3.26)
      else
        . Use current solution updates $G_j$ and $G_i^*$ and $\tilde{\alpha}_n^\lambda$
        to compute $B_{\text{inv}}^m$ in (3.26)
      end
      if $h_m > B_{\text{inv}}^m$
        . $M_e \leftarrow M_e \bigcup \{\Delta_m\}$
      end
    end
    if $M_e \neq \{\}$
      . Refine the marked elementsand update themesh $\Delta$
      . $M_e \leftarrow \{\}$
    else
      . flag = False
    end
  end
◇ Solve for $\tilde{\alpha}_n^\lambda$.
```

Remark 2.

(i) In practice, B_{inv}^m in (3.26) cannot be computed since α^λ, g_j, g_i^* and $\mathcal{T}_n$ are unknown. Similar to the approach described in section 3.1, we can compute an estimate for B_{inv}^m by using the uniform boundedness of the operator $\mathcal{T}_n$ [9] and by using approximate values for the functions involved in B_{inv}^m. In this context, we use either *a priori* information or the recent forward and inverse problem solution updates to calculate (3.26) on each element. Then, the elements with the mesh parameter $h_m > B_{\text{inv}}^m$ are determined and refined.

(ii) In order to save computations, after the first sweep of refinement, one can compute the bound B_{inv}^m only for the new elements. Furthermore, similar to the approach described in section 3.1, the term $\|g_i^* g_j\|_{0,m}$ in (3.26) can be stored after the first mesh generation and can be used in the following mesh generations. In this context, the bound B_{inv}^m can be updated by using only the updated $\|\alpha^\lambda\|_{1,m}$ value.

(iii) Note that, in practice, one of the two terms in the denominator of B^m_{inv} will be dominant depending on the value of λ. Thus, we consider only the dominant term for the computation of B^m_{inv}. In case ϵ_{inv} cannot be chosen in prior, we consider a posterior approach, set $\epsilon_{\text{inv}} = 1$ and compute $\|\alpha^\lambda\|_{1,m} h_m$ or $\sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m} \|\alpha^\lambda\|_{1,m} h_m$ on each element, which are used as the refinement indicators. Then, the elements with indicator value which exceeds the average indicator value are refined. In this case, the algorithm has to be stopped when the number of nodes in the mesh exceeds the allowable number of nodes.

3.4. Computational cost of the adaptive mesh generation algorithm for the inverse problem

Consider the algorithm stated in remark 2(iii) for $\Omega \subset \mathbb{R}^2$ and assume that the second term in the denominator of B^m_{inv} (3.26) is dominant. Using triangular finite elements with first-order Lagrange basis functions and an analytical (exact) integration on each finite element, the total number of multiplications required to compute the error estimates on all finite elements is given by $(120 N_d N_s + 14) N_\Delta$.

In order to reduce the number of multiplications, we can consider an approximation for B^m_{inv} as follows:

$$B^m_{\text{inv}} \approx \frac{1}{\left\| \sum_{i,j}^{N_d,N_s} g_i^* g_j \right\|_{0,m} \|\alpha^\lambda\|_{1,m}}. \tag{3.28}$$

Then, the number of multiplications reduces to $134 N_\Delta$.

If one uses the interpolation error estimate (2.14) to generate adaptive meshes, the number of multiplications to compute the error estimates on all finite elements will be $13 N_\Delta$. However, such adaptive meshes may not help reduce the error in the reconstructed optical images, resulting from discretization (see simulation experiment 3).

4. Numerical experiments

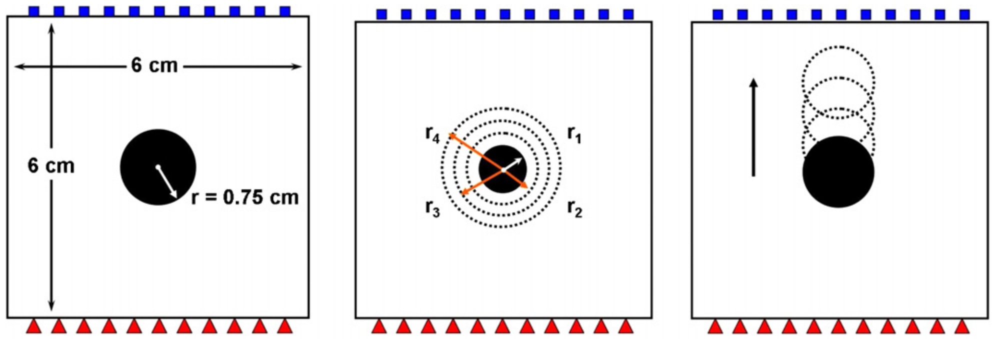
We conduct a series of numerical experiments to demonstrate the implications of theorems 1 and 2, and to present the effectiveness of the proposed adaptive mesh generation algorithms. We perform our experiments in 2D for ease of comparison.

In the first simulation, we consider a series of image reconstructions to show the effectiveness of the proposed adaptive mesh generation algorithms. In this context, we compare the images reconstructed by using uniform meshes for the forward and inverse problems to the images reconstructed by using adaptive meshes which are designed based on theorems 1 and 2.

In the second simulation, we show the effect of the heterogeneity size on the accuracy of the reconstructed absorption images. Next, we demonstrate how this error can be addressed by the proposed adaptive discretization schemes.

In the final simulation study, we demonstrate the implication of theorem 2 and show that meshes generated for the forward problem by using discretization error estimates which disregard the interaction between the solutions g_j, g_i^* and α^λ can lead to unstable image reconstructions. We note that the proposed adaptive mesh generation algorithm for the forward problem addresses this problem.

Note that in all experiments, we use triangular finite elements with first-order Lagrange basis functions. We apply Gaussian elimination method to solve the discrete forward problem resulting from the variational formulation [5] of the boundary value problems (2.1)–(2.2)



(a) (b) (c)

Figure 1. The setups used for the simulation studies 1, 2 and 3. The squares and triangles denote the detectors and sources, respectively. (a) The optical domain and source–detector configuration for simulation study 1. (b) The optical domain and source–detector configuration for simulation study 2. $r_1 = 0.50$ cm, $r_2 = 0.75$ cm, $r_3 = 1.0$ cm and $r_4 = 1.25$ cm. (c) The optical domain and source–detector configuration for simulation study 3. The radius of the circles is 0.75 cm.

and (2.3)–(2.4) [9]. For the inverse problem, we consider the discrete problem obtained by projecting (2.18) by the collocation method [9]:

$$\mathcal{P}_n \tilde{\mathcal{K}} \tilde{\alpha}_n^{\lambda} = \mathcal{P}_n \tilde{\gamma}, \tag{4.29}$$

where the regularization parameter is chosen as small as possible, yet large enough to enable robust image reconstructions. In this respect, an appropriate value for the regularization parameter is chosen based on experience. The discrete inverse problem (4.29) is solved using Gaussian elimination as well.

4.1. Simulation study 1

In this simulation study, we consider the geometry shown in figure 1(a). We simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 81 nodes along the x and y directions, where the refractive index mismatch parameter $a = 3.11$ sources and 11 detectors are evenly spaced on the bottom and top edges of the square, respectively. The diffusion coefficient $D(\mathbf{x}) = 0.0410$ for $\mathbf{x} \in \Omega \cup \partial\Omega$. The circular heterogeneity with absorption coefficient $\mu_a = 0.2\ \mathrm{cm}^{-1}$ is embedded in an optically homogeneous background with $\mu_a = 0.04\ \mathrm{cm}^{-1}$.

In order to obtain a series of absorption imaging problems using the same setup, we consider five values for the background absorption value. Then, for each imaging problem, we consider three mesh scenarios: uniform mesh for both forward and inverse problems; adaptive mesh for the forward problem and uniform mesh for the inverse problem; and adaptive meshes for both forward and inverse problems. We refer to table 2 for a brief outline of the first simulation study.

The uniform mesh used for the forward problem discretization has 625 nodes and is shown in figure 2(a). The uniform mesh for the inverse problem has 313 nodes and is shown in figure 2(b). We use the algorithms described in section 3.1 and remark 1(iii), and section 3.3 remark 2(iii) to generate the adaptive meshes for the forward and inverse problems, respectively. The number of nodes in each of the adaptive meshes used for the forward problem does not exceed 750. An example for the adaptive mesh generated for a source located at (1.0, 0) is shown in figure 2(c). The adaptive mesh for the inverse problem generated for the case where the background $\mu_a = 0.050\ \mathrm{cm}^{-1}$ has 418 nodes and is shown in figure 2(d).

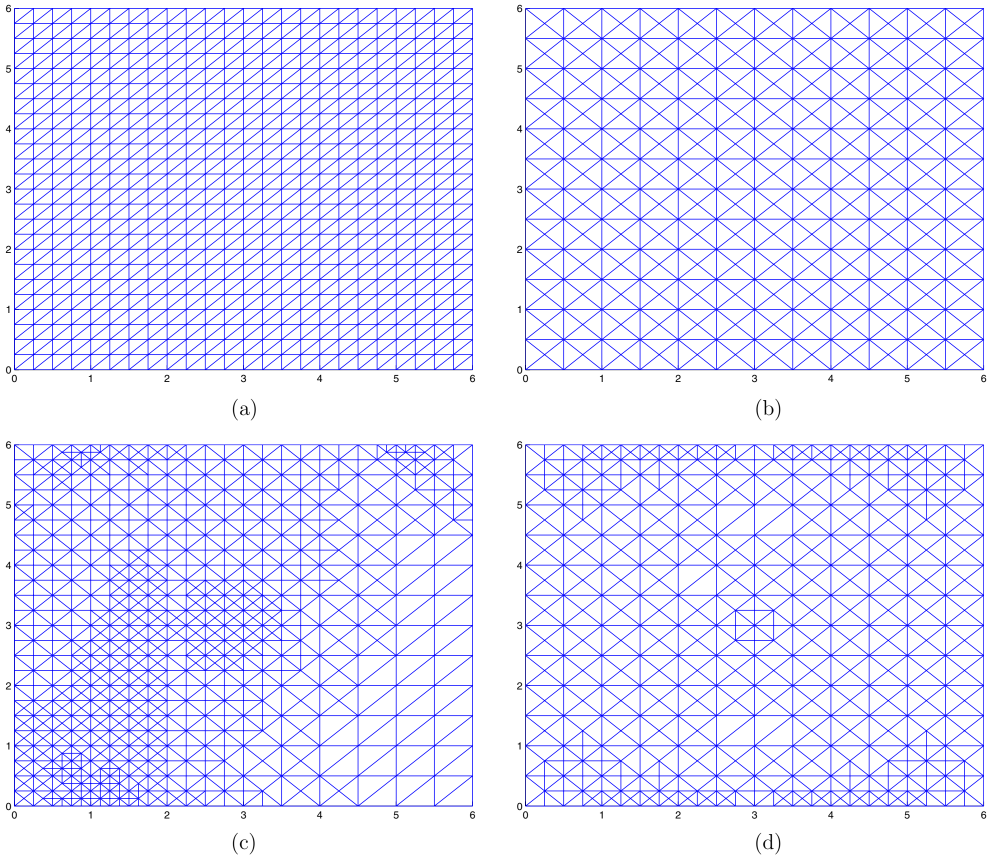
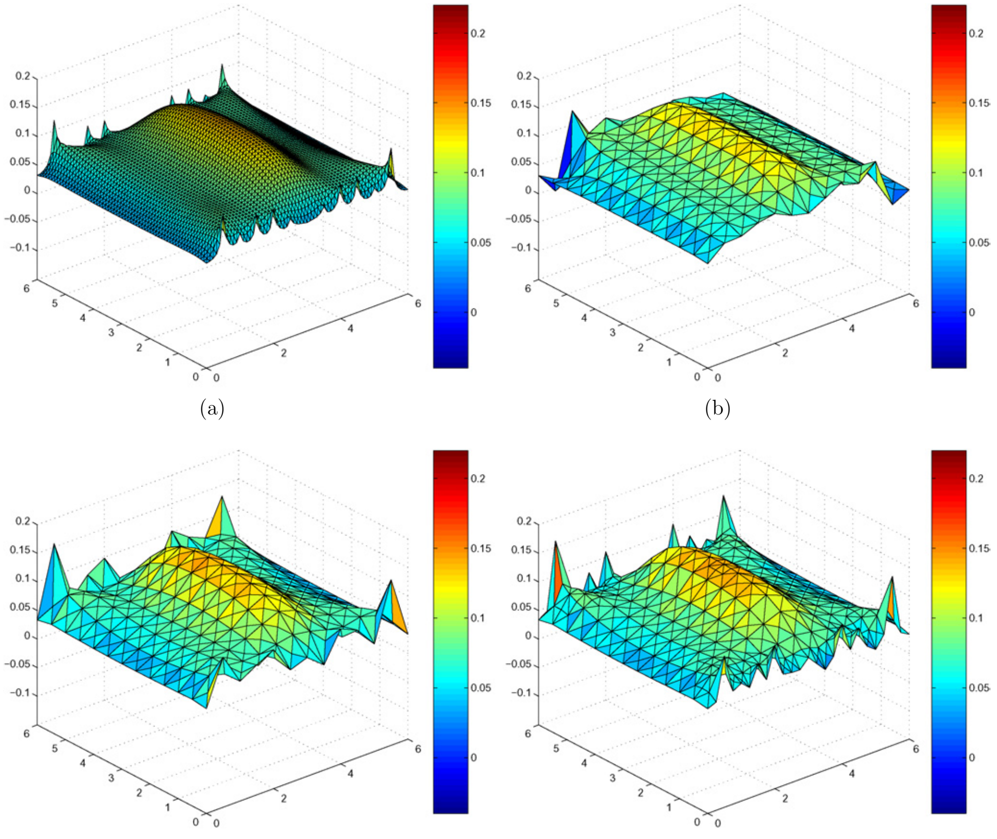

(a) (b) (c) (d)

Figure 2. Examples of meshes used in the first simulation study. (a) The uniform mesh with 625 nodes. (b) The uniform mesh with 313 nodes. (c) The adaptive mesh generated for the forward problem for the source located at (1.0,0): background $\mu_a = 0.050\ \text{cm}^{-1}$. (d) The adaptive mesh generated for the inverse problem solution, with 418 nodes. Background $\mu_a = 0.050\ \text{cm}^{-1}$.

Table 2. The mesh scenarios and the background μ_a values in simulation study 1.

Mesh (forward)	Mesh (inverse)	Background μ_a (cm^{-1})
Uniform	Uniform	0.032, 0.036, 0.040, 0.044, 0.050
Adaptive	Uniform	0.032, 0.036, 0.040, 0.044, 0.050
Adaptive	Adaptive	0.032, 0.036, 0.040, 0.044, 0.050

For the inverse problem, we set the regularization parameter λ to 10^{-7} in all experiments to eliminate the dependence of the error estimates (2.15)–(2.19) on the regularization parameter. We consider the image reconstructed by using fine uniform meshes (61×61 nodes for the forward problem and 61×61 nodes for the inverse problem) as the reference image α^{λ}, which is assumed to possess no error due to discretization. We compute the error $\left\|\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda}\right\|_{L^1(\Omega)}$ for each image reconstruction and tabulate the results in table 3. We see that the error in the images reconstructed by using uniform meshes for both forward and inverse problems is significantly reduced by the use of adaptively refined meshes. A similar behaviour is observed for all choices of background absorption value.



(a) (b)

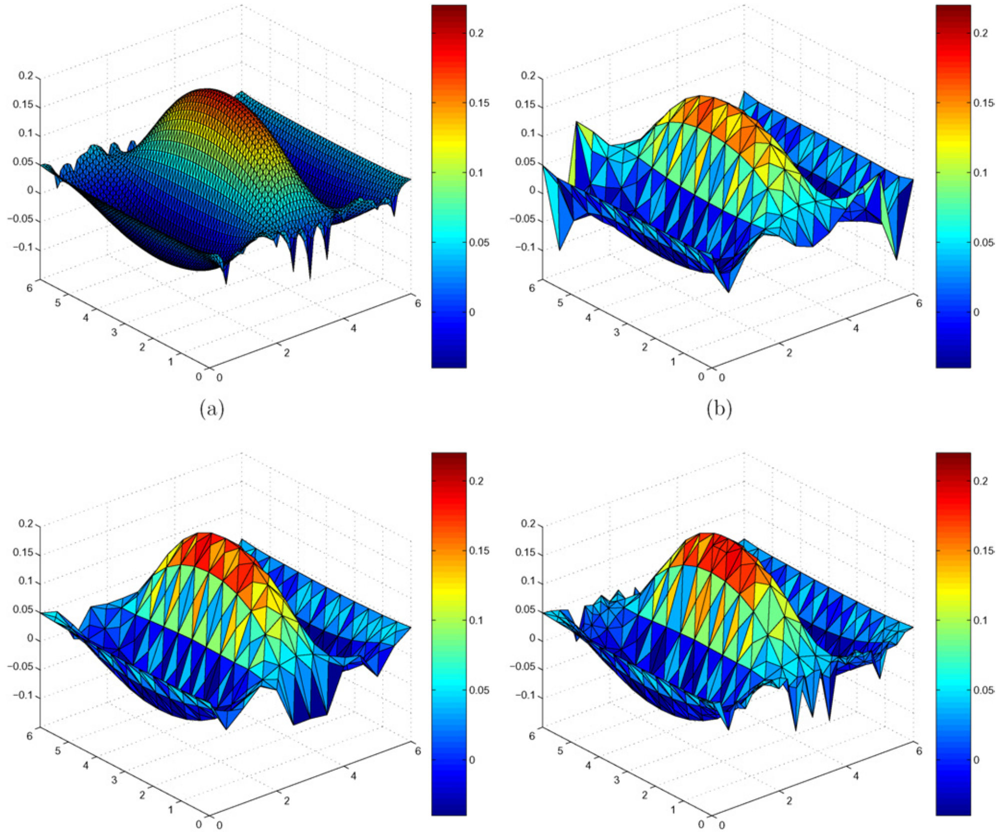
(c) (d)

Figure 3. The reconstruction results of simulation study 1, with the background $\mu_a = 0.032\ \mathrm{cm}^{-1}$. (a) The optical absorption image used as the reference for error computations. (b) The reconstructed absorption image using the uniform mesh in figure 2(a) for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (c) The reconstructed absorption image using an adaptive mesh for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (d) The reconstructed absorption image using an adaptive mesh for the forward, and the adaptive mesh in figure 2(d) for the inverse problem.

Table 3. The error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 1 and table 2. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The unit of background μ_a is cm^{-1}.

	Background μ_a:	0.032	0.036	0.040	0.044	0.050
Uniform–uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2325	0.2559	0.2773	0.2932	0.3013
Adaptive–uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.1238	0.1139	0.1166	0.1209	0.1278
Adaptive–adaptive	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.1043	0.0997	0.0998	0.1003	0.1009

We present image reconstructions in figures 3 and 4 for the two extreme cases, where the background absorption value is equal to 0.032 and 0.050 cm^{-1}, respectively. Figures 3(a) and 4(a) display the reference images used to compute the error values given in table 3.



(a) (b)

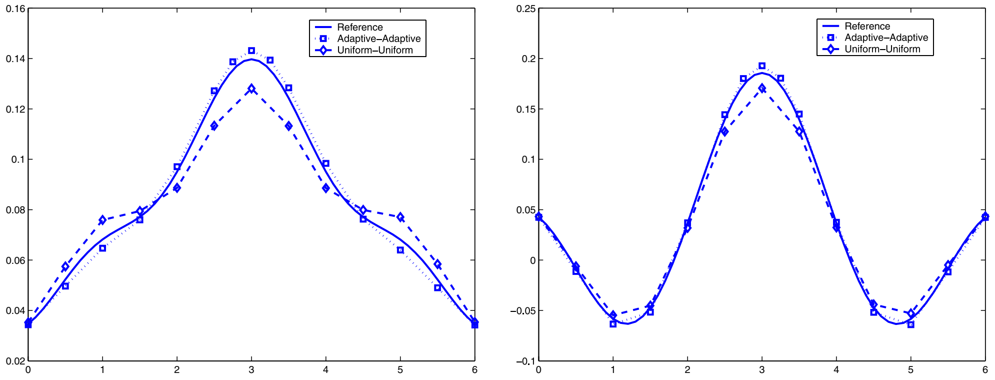
(c) (d)

Figure 4. The results of simulation study 1, with the background $\mu_a = 0.050\ \mathrm{cm}^{-1}$. (a) The optical absorption image used as the reference for error computations. (b) The reconstructed absorption image using the uniform mesh in figure 2(a) for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (c) The reconstructed absorption image using an adaptive mesh for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (d) The reconstructed absorption image using an adaptive mesh for the forward, and the adaptive mesh in figure 2(d) for the inverse problem.

Figures 3(c) and (d) show that the optical heterogeneity is resolved better by using adaptive meshes as compared to the reconstructed image obtained by using uniform meshes, which is shown in figure 3(b). These results are consistent with the error values given in table 3. A similar trend is seen in figures 4(c) and (d). Note that the number of nodes in the adaptive meshes is almost equal to the number of nodes that the uniform meshes have. In figure 5, we show the cross-sectional views from the reconstructed images. We see that the use of coarse uniform meshes fails to resolve the circular heterogeneity especially for the case in which the background $\mu_a = 0.032\ \mathrm{cm}^{-1}$.

4.2. Simulation study 2

In this study, we consider the geometry shown in figure 1(b). To simulate the optical data, we use the same source–detector configuration considered in the first simulation study. We



(a) (b)

Figure 5. The cross-sectional views from the reconstructed images in simulation study 1, corresponding to the cases where the background $\mu_a = 0.032$ and $\mu_a = 0.050\ \text{cm}^{-1}$, respectively. (a) The cross-sectional cuts taken from figures 3(a), (b) and (d), along the x-direction at $y = 3$. The solid, square and diamond lines correspond to the cross-sectional cuts taken from the images shown in figures 3(a), (b) and (d), respectively. (b) The cross-sectional cuts taken from figures 4(a), (b) and (d), along the x-direction at $y = 3$. The solid, square and diamond lines correspond to the cross-sectional cuts taken from the images shown in figures 4(a), (b) and (d), respectively.

simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 81 nodes along the x and y directions, where the refractive index mismatch parameter $a = 3$. The diffusion coefficient D is assumed to be constant and $D(\mathbf{x}) = 0.0410$ cm, for all $\mathbf{x} \in \Omega \cup \partial\Omega$.

We consider four different radii for the circular heterogeneity with $\mu_a = 0.20\ \text{cm}^{-1}$ embedded in a background with $\mu_a = 0.040\ \text{cm}^{-1}$ as shown in figure 1(b). For each case, we compute the error for different mesh scenarios, similar to the first simulation study: uniform mesh for both forward and inverse problems; adaptive mesh for the forward problem and uniform mesh for the inverse problem; and adaptive meshes for both forward and inverse problems. The adaptive meshes for this simulation study were generated based on theorems 1 and 2, and the mesh generation algorithms described in the first simulation study and section 3. The uniform meshes used for the forward and inverse problems are identical to those used in the first simulation study. We note that the number of nodes in the adaptive meshes generated for the forward and inverse problems is close to the number of nodes in the corresponding uniform meshes.

In table 4, we tabulate the error norm $\left\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\right\|$ obtained for each heterogeneity size with different mesh choices, where α^λ is the reference image reconstructed by using fine uniform meshes as in the first study. Table 4 shows that the error increases with increasing heterogeneity size. We see that the reduction in the error as a result of using adaptive meshes is more significant for smaller sized heterogeneities. Further reduction in the error norm $\left\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\right\|$ is possible by increasing the number of nodes in the meshes.

For brevity, we only show the reconstruction results for the extreme cases: $r = 0.5$ cm and $r = 1.25$ cm. We note that the regularization parameter $\lambda = 5 \times 10^{-9}$ in all reconstructions. Figures 6(a) and (b) show the images used as the reference images α^λ in the calculation of the error norms $\left\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\right\|$ listed in table 4. Figures 6(e)–(f) show that the adaptive meshes reduce the artefacts as compared to the images reconstructed by using uniform meshes, which are shown in figures 6(c)–(d).

Table 4. The L^1 norm of α^λ and the error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 2. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The radius of the circular heterogeneity is given in cm.

	Radius:	0.50	0.75	1.0	1.25
	$\|\alpha^\lambda\|_{L^1(\Omega)}$:	0.7196	1.3760	1.4759	1.7817
Uniform–uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.5622	0.5706	0.5850	0.6337
Adaptive–uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2153	0.2776	0.3766	0.5113
Adaptive–adaptive	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2020	0.2630	0.3592	0.5034

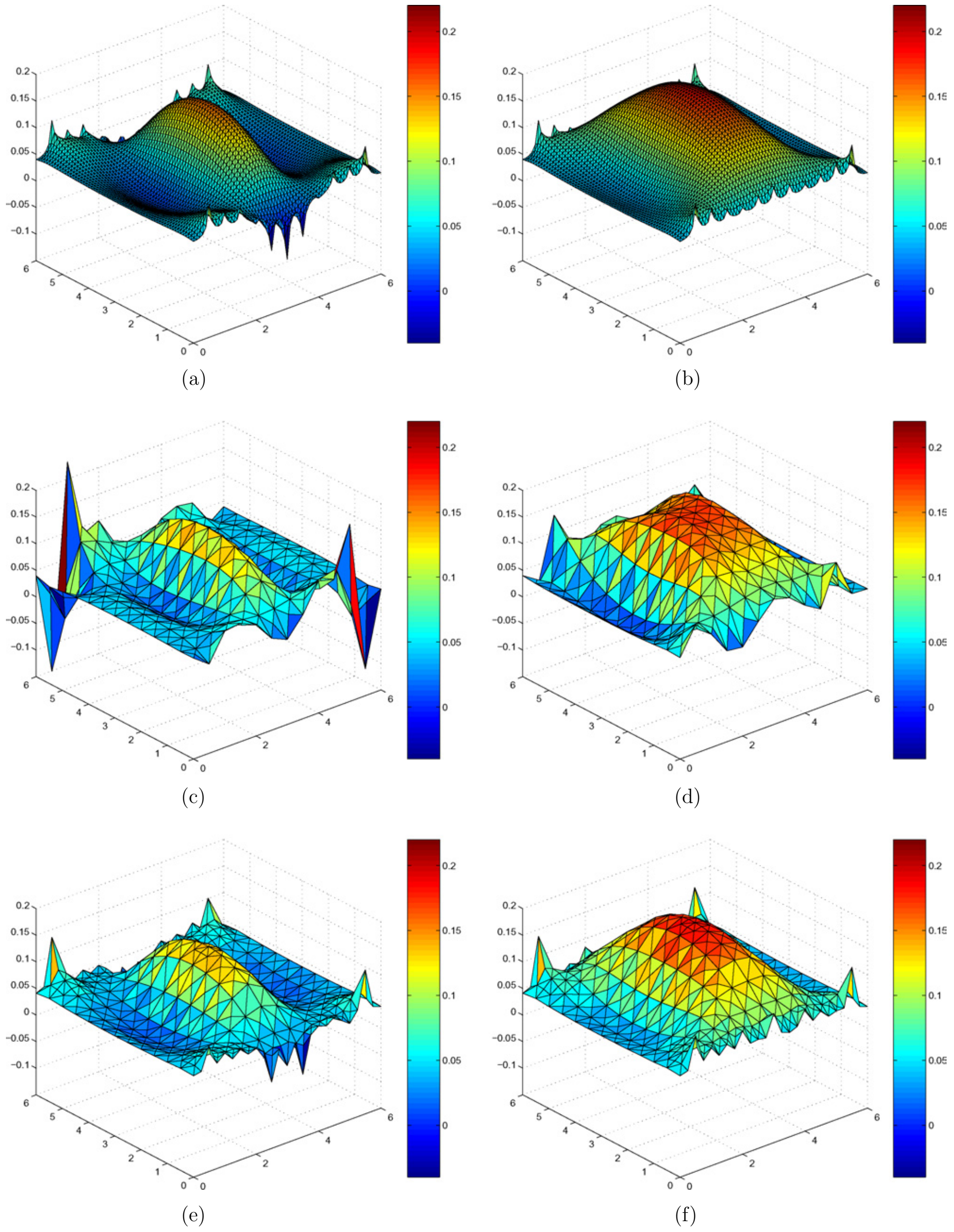
4.3. Simulation study 3

In this simulation study, we consider the geometry shown in figure 1(c). The centre of the circular heterogeneity is moved vertically towards the detector side to see the effect on the imaging accuracy. Next, we show how the error in the reconstructed images due to discretization can be addressed by using appropriate meshes for the solutions of the forward and inverse problems. In this context, we compare the results obtained by using (1) uniform meshes, (2) the adaptive meshes generated using conventional *a priori* discretization error estimates, and (3) the adaptive meshes proposed in this study. By conventional error estimates, we mean the *a priori* discretization error estimates (2.16) and (2.17) for the solution of the forward problem, and the *a priori* interpolation error estimate (2.14) for the solution of the inverse problem.

To simulate the optical data, we use the same source–detector configuration considered in the first simulation study. We simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 81 nodes along the x and y directions, where the refractive index mismatch parameter $a = 3$. We note that, in all reconstructions, the background absorption value is set to $\mu_a = 0.04$ cm^{-1} and the diffusion coefficient D is assumed to be constant and $D(\mathbf{x}) = 0.0410$ cm, for all $\mathbf{x} \in \Omega \cup \partial\Omega$.

The uniform meshes used in this simulation study are identical to those used in simulation studies 1 and 2. Sample meshes for the forward problem solution using the conventional and the proposed adaptive meshing strategies are shown in figures 7(a) and (b) and figures 8(c) and (d), respectively. We see that the conventional adaptive mesh generation strategy leads to meshes refined around only sources or detectors, but not both. In contrast, figures 8(c) and (d) show that the proposed strategy results in adaptive meshes refined around sources, detectors and the heterogeneity as well. This observation is consistent with theorem 2. The adaptive mesh for the inverse problem solution, which was generated using the *a priori* interpolation error estimate (2.14), is shown in figure 8(f). Note that the mesh was generated for the case where the circular heterogeneity was centred at (3.0, 3.5). The mesh generated based on theorem 1 (figure 8(e)) provides higher resolution close to the sources and detectors as compared to the mesh shown in figure 8(f), which is merely refined around the heterogeneity.

In this simulation study, we consider four different positions for the centre of the circular heterogeneity with radius 0.75 cm, along the y-axis: centre at (3.0, 3.0), (3.0, 3.5), (3.0, 4.0) and (3.0, 4.5), respectively. Similar to the previous simulations, we compute the error in the reconstructed images for all cases, and compare the error values attained by different meshing strategies. Finally we present the reconstructed images obtained by using different mesh strategies corresponding to the case where the circular inclusion is centred at (3.0, 3.5) and (3.0, 4.0).

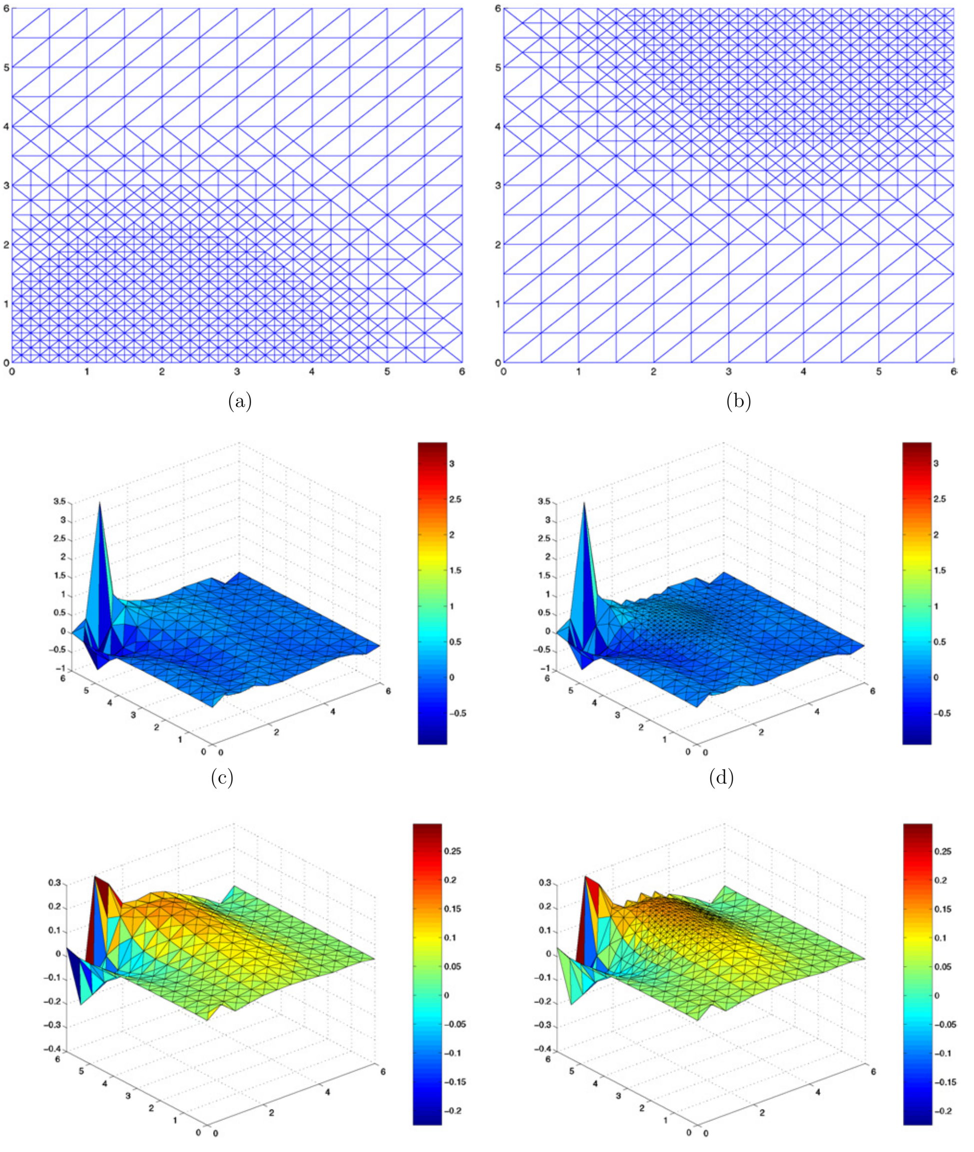


(a) (b)

(c) (d)

(e) (f)

Figure 6. The results of simulation study 2. The left and right columns show the reconstructed images regarding the optical heterogeneity with radius 0.50 cm and 1.25 cm, respectively. The background $\mu_a = 0.040$ cm^{-1} in all of the reconstructions. The reference images shown in (a) and (b) are obtained using a uniform mesh with 61×61 nodes in both the forward and inverse problems. ((a) and (b)) The optical absorption images used as the reference for error computations. The images correspond to the reconstruction of the circular heterogeneities of radii 0.5 cm and 1.25 cm, respectively. ((c) and (d)) The reconstructed absorption images using the uniform mesh in figure 2(a) for the forward, and the uniform mesh in figure 2(b) for the inverse problem. ((e) and (f)) The reconstructed absorption image using adaptive meshes for both the forward and the inverse problems.



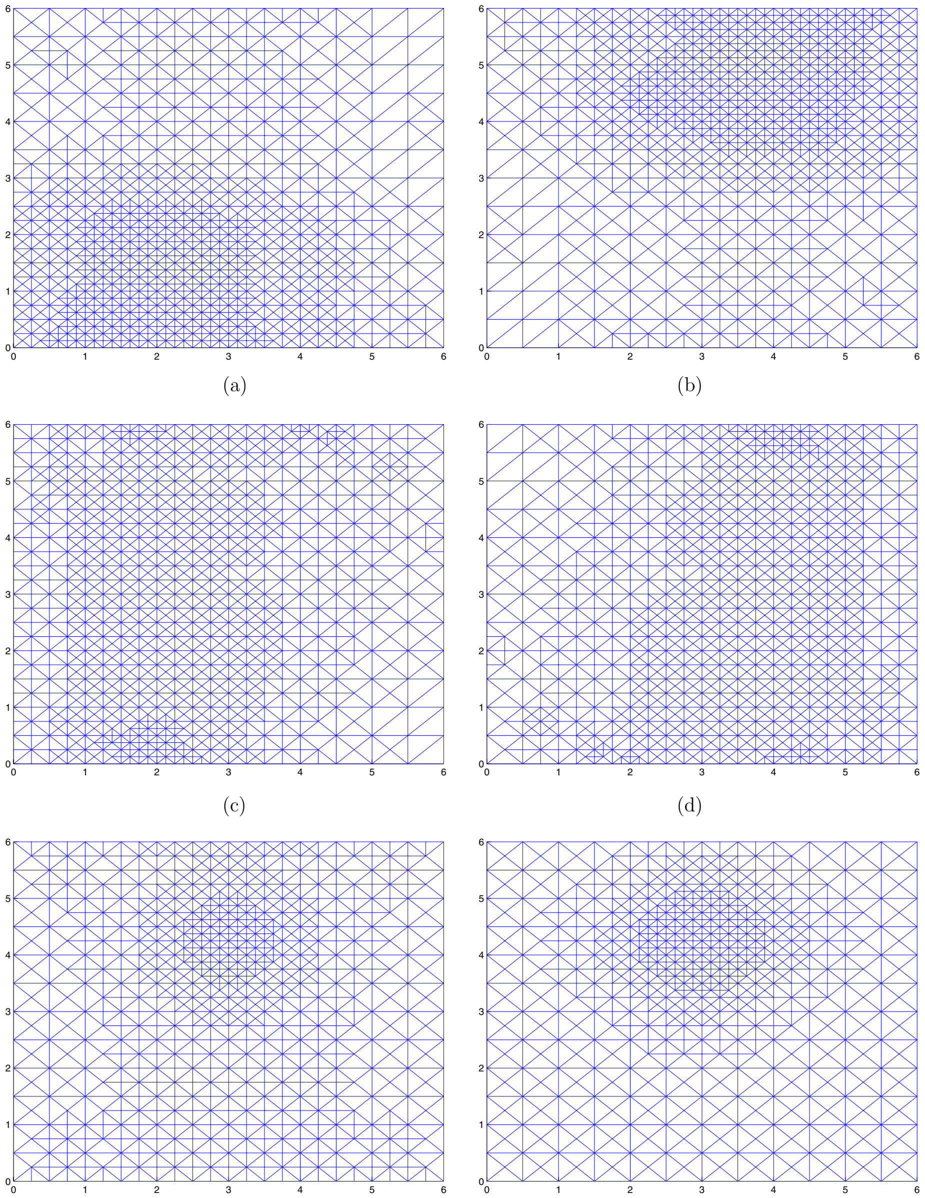
(a) (b)

(c) (d)

(e) (f)

Figure 7. ((a) and (b)) Samples of adaptive meshes in the third simulation study (with 865 nodes for the source and the detector located at (2.0, 0) and (4.0, 6.0), respectively), generated by using the conventional error estimates (2.16) and (2.17), which led to unstable optical image reconstruction shown in (c) to (f), for the circular heterogeneity centred at (3.0, 3.5). ((c) and (d)) The unstable optical image reconstructions in the third simulation study, obtained by using the adaptive meshes for the forward problem solution whose examples are shown in (a) and (b) ($\lambda = 10^{-8}$). ((e) and (f)) The unstable optical image reconstructions in the third simulation study, obtained by using the adaptive meshes for the forward problem solution whose examples are shown in (a) and (b). λ was set to 10^{-6} to suppress the significantly large artefacts observed in (c) and (d).

Using the meshes for the forward problem discretization (see figures 7(a) and (b)), which were generated by using the conventional *a priori* discretization error estimates (2.16)–(2.17), leads to the image reconstructions shown in figures 7(c) and (d), where the



(a) (b)

(c) (d)

(e) (f)

Figure 8. Samples of adaptive meshes used in the third simulation study, which led to the optical image reconstructions shown in figure 10. The meshes were generated for the circular heterogeneity centred at (3.0, 4.0). (a) The adaptive mesh with 942 nodes for the forward problem solution for the source located at (2.0, 0), obtained by refining the adaptive mesh shown in figure 7(a) around the detectors. (b) The adaptive mesh with 955 nodes for the forward problem solution for the detector located at (4.0, 6.0), obtained by refining the adaptive mesh shown in figure 7(b) around the sources. (c) The adaptive mesh with 895 nodes for the forward problem solution for the source located at (2.0, 0), generated based on theorem 2. (d) The adaptive mesh with 896 nodes for the forward problem solution for the detector located at (4.0, 6.0), generated based on theorem 2. (e) The adaptive mesh with 691 nodes for the inverse problem solution, generated based on theorem 1. (f) The adaptive mesh with 609 nodes for the inverse problem solution, generated based on the conventional error estimate (2.14).

regularization parameter $\lambda = 10^{-8}$. We observe that the finite-dimensional operator does not provide a stable solution. We note that using an adaptive mesh for the inverse problem solution does not change the outcome (figure 7(d)). Note also that the meshes generated by using the conventional *a priori* discretization error estimates (2.16)–(2.17) are sufficient to provide accurate finite element approximations to the actual solutions g_j and g_i^*. Therefore, the unstable reconstructions can be attributed to the errors $\mathcal{K} - \tilde{\mathcal{K}}$ and $\gamma - \tilde{\gamma}$, due to inappropriate discretization as noted by theorem 2. In consistence with theorem 2, this observation suggests that solving the forward problem accurately does not necessarily imply that approximate operator $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ are error-free. Therefore, in order to address such problems, one has to follow a discretization scheme based on theorem 2 for the solution of the forward problem, which takes into account the interaction between the solutions of the diffusion equation and the associated adjoint problem, as described in section 3.1.

In order to suppress the severe artefacts observed in figures 7(c) and (d), we increased the regularization parameter and set $\lambda = 10^{-6}$. The resulting images are shown in figures 7(e) and (f). As noted by theorems 1 and 2, increasing the regularization parameter reduces the error in the reconstructed images. However, increasing the regularization parameter will also compromise the image quality and lead to over-smoothed images. In order to address the instability issue without degrading the image quality by using high regularization parameters, we modified the adaptive mesh generation method that leads to the meshes shown in figures 7(a) and (b). In this context, for the first 2 refinements, we used the proposed mesh generation algorithm based on theorem 2 to generate an initial adaptive mesh; and for the next two refinements, we used the conventional error estimates (2.16)–(2.17). Following this modification, the samples of the resulting adaptive meshes are shown in figures 8(a) and (b). For a comparison, we also present in figures 8(c) and (d), the adaptive meshes generated by using the proposed adaptive mesh generation algorithms as described in section 3.1 and remark 1(iii). We observe that the meshes shown in figures 8(c) and (d), indicate further refinement around sources, detectors and the circular heterogeneity as compared to the adaptive meshes shown in figures 8(a) and (b).

Examples of the adaptive meshes generated for the inverse problem based on theorem 1 and the conventional *a priori* interpolation estimate (2.14) are shown in figures 8(e) and (f), respectively. We observe that the adaptive mesh shown in figure 8(e) provides higher resolution around sources and detectors as compared to the adaptive mesh shown in figure 8(f).

We note that the uniform meshes used in this simulation study are identical to those used in the previous simulation studies.

In order to compare the performance of the conventional and proposed adaptive mesh strategies, we perform four experiments and compute the error in the reconstructed optical absorption images. For each experiment, we consider five different mesh strategies and refer to table 5 for the description of these experiments.

We show the reconstructed optical absorption images for the two cases in figures 9 and 10, corresponding to the circular heterogeneity centred at (3.0, 3.5) and (3.0, 4.0), respectively. Figures 9(a) and 10(a) show the reference absorption image reconstructions which are used to compute the error in the reconstructed optical images.

Figure 9(b) shows the image reconstructed using coarse uniform meshes for both the forward and inverse problems, for the case where the circular inclusion is centred at (3.0, 3.5) where the regularization parameter was set to $\lambda = 10^{-8}$. With the same value of the regularization parameter, figure 9(c) shows the reconstructed image by using the adaptive mesh based on theorem 2 for the forward problem and the coarse uniform mesh (shown in figure 2(b)) for the inverse problem. Figure 9(e) shows the reconstructed image obtained

Table 5. The relevant parameters in experiments 1–5 in simulation study 3. The abbreviation ‘Conv.’ implies that the corresponding mesh was generated using the conventional *a priori* discretization error estimates (2.16)–(2.17) for the forward problem solution, and the conventional *a priori* interpolation error estimate (2.14) for the inverse problem solution. The abbreviation ‘Prop.’ refers to the adaptive meshes generated by using the proposed adaptive mesh generation algorithms based on theorems 1 and 2, for the inverse and forward problem solutions, respectively. The last column in the table shows the coordinates of the centre of the circular heterogeneity, considered in each experiment.

	Mesh (forward)	Mesh (inverse)	Centre at:
Exp. 1	Uniform	Uniform	[(3.0, 3.0), (3.0, 3.5), (3.0, 4.0), (3.0, 4.5)]
Exp. 2	Adaptive (Conv.)	Uniform	[(3.0, 3.0), (3.0, 3.5), (3.0, 4.0), (3.0, 4.5)]
Exp. 3	Adaptive (Conv.)	Adaptive (Conv.)	[(3.0, 3.0), (3.0, 3.5), (3.0, 4.0), (3.0, 4.5)]
Exp. 4	Adaptive (Prop.)	Uniform	[(3.0, 3.0), (3.0, 3.5), (3.0, 4.0), (3.0, 4.5)]
Exp. 5	Adaptive (Prop.)	Adaptive (Prop.)	[(3.0, 3.0), (3.0, 3.5), (3.0, 4.0), (3.0, 4.5)]

Table 6. The error $\|\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda}\|_{L^1(\Omega)}$ for each experiment described in the simulation study 3. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The superscript ‘C’ denotes that the corresponding adaptive mesh generation is based on the conventional *a priori* error estimates (2.16)–(2.17) and (2.14).

	Radius at:	(3.0, 3.0)	(3.0, 3.5)	(3.0, 4.0)	(3.0, 4.5)
Uniform–uniform	$\|\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda}\|_{L^1(\Omega)}$:	0.4539	0.4606	0.4733	0.4956
Adaptive–uniform	$\|\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda}\|_{L^1(\Omega)}$:	0.2690	0.2695	0.2634	0.2507
Adaptive–adaptive	$\|\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda}\|_{L^1(\Omega)}$:	0.2433	0.2455	0.2459	0.2434
Adaptive–uniform	$\|\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda,C}\|_{L^1(\Omega)}$:	0.7989	0.7596	0.7072	0.6418
Adaptive–adaptive	$\|\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda,C}\|_{L^1(\Omega)}$:	0.8011	0.7614	0.7070	0.6351

by using the adaptive meshes based on theorems 1 and 2. We observe the improvements especially around the boundaries. Using the conventional adaptive meshes for the forward problem solution, which were modified around sources and detectors as noted before, we ran into a similar instability problem. Therefore, in order to obtain better reconstructions with the conventional adaptive meshes, we set the regularization parameter $\lambda = 10^{-7}$ in the corresponding inverse problem formulations. The resulting reconstructed images are shown in figures 9(c) and (f). In this case, we observe that the use of conventional adaptive meshes for the forward and inverse problems does not improve the image quality as compared to the reconstructed image shown in figure 9(b), which is obtained by using coarse uniform meshes.

We observe similar results for the case where the circular inclusion is centred at (3.0, 4.0). We note that the regularization parameter is set to $\lambda = 10^{-8}$ for all reconstructions except for the reconstructions obtained by using conventional adaptive meshes, in which case $\lambda = 10^{-7}$. Figure 9 shows the reconstructed images corresponding to all meshing strategies.

Table 6 shows the error norm $\alpha^{\lambda} - \tilde{\alpha}_n^{\lambda}$ computations for all cases. Similar to the previous experiments, α^{λ} is the reference image obtained by using fine uniform meshes for the discretization of the forward and inverse problems. The error values are consistent with figures 9 and 10. In all cases, the proposed adaptive meshes significantly reduce the error in the reconstructed images. Furthermore, the image quality is enhanced by merely appropriate discretization, without having to increase the regularization parameter. In contrast, the conventional adaptive meshes perform worse than uniform meshes even though a higher regularization parameter is used.

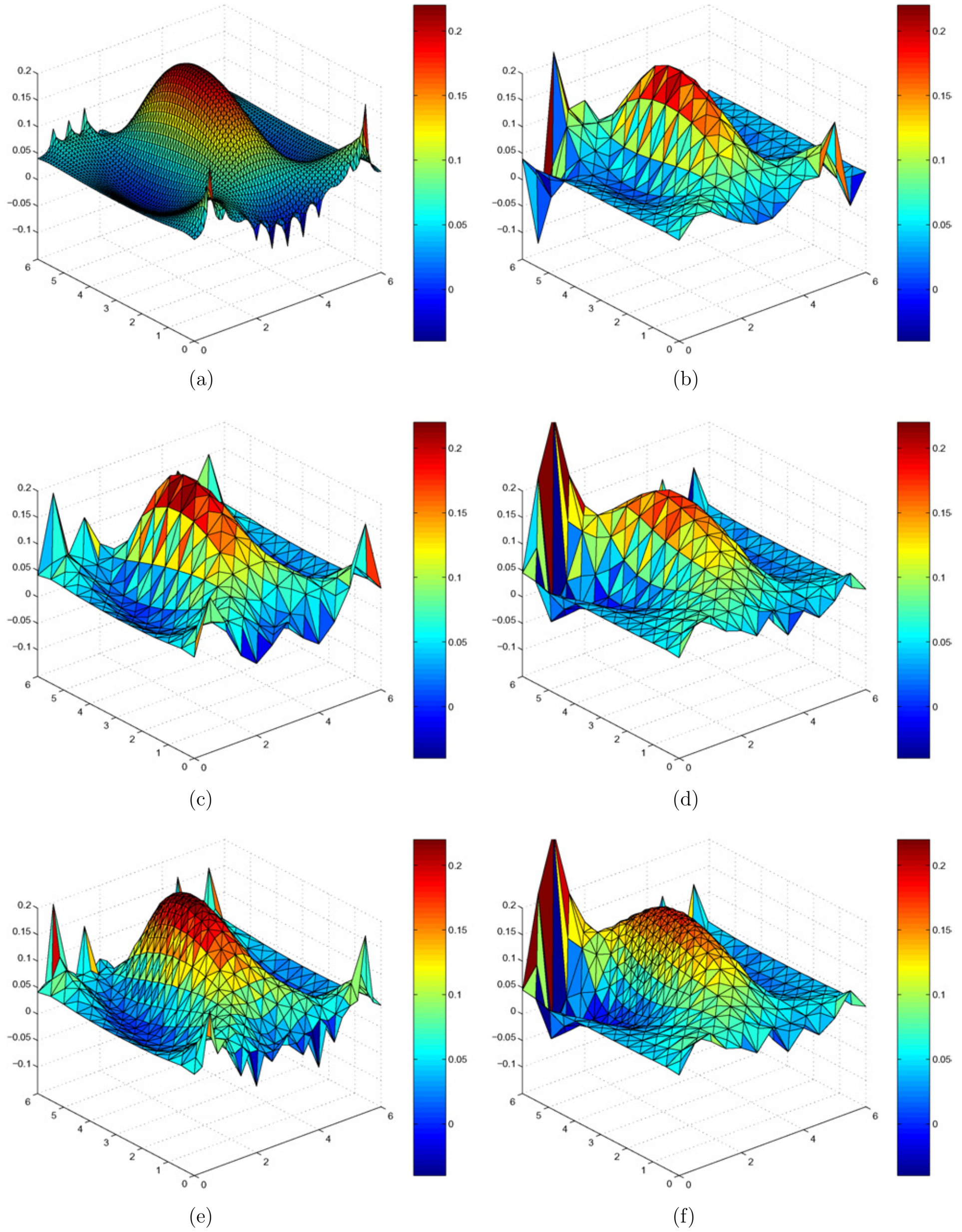
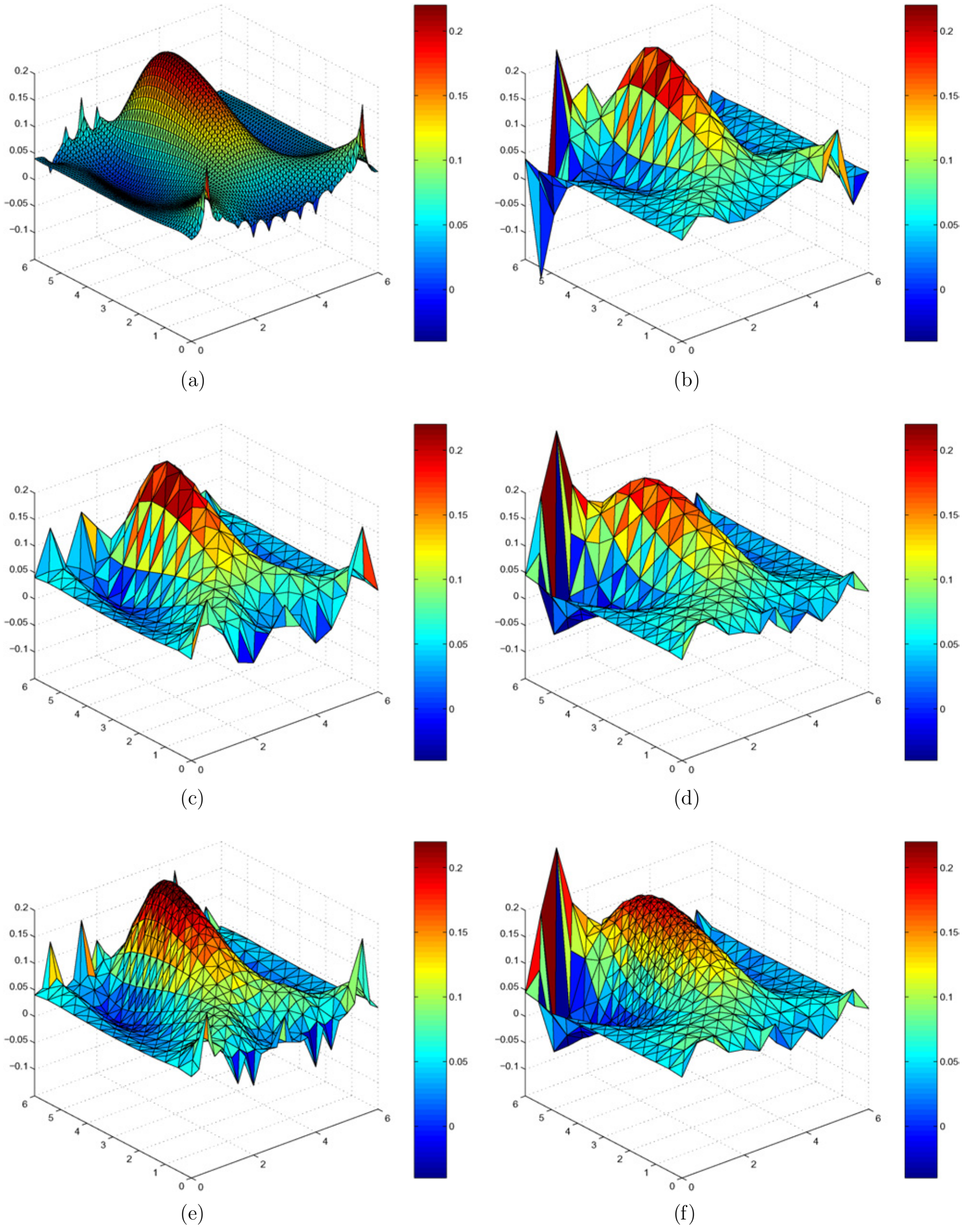


Figure 9. The reconstructed optical images regarding the circular heterogeneity centred at (3.0, 3.5) in the third simulation study. (a) The absorption image used as the reference in the error computations. (b) The reconstructed absorption image using the uniform mesh in figure 2(a) for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (c) The reconstructed absorption image using adaptive meshes based on theorem 2 for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (d) The reconstructed absorption image using adaptive meshes based on *a priori* error estimates (2.16) and (2.17) for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (e) The reconstructed absorption image using adaptive meshes based on theorem 2 for the forward, and using the adaptive mesh based on theorem 1 for the inverse problem. (f) The reconstructed absorption image using adaptive meshes based on *a priori* error estimates (2.16) and (2.17) for the forward, and the interpolation error estimate (2.14) for the inverse problem.



(a) (b)

(c) (d)

(e) (f)

Figure 10. The reconstructed optical images regarding the circular heterogeneity centred at $(3.0, 4.0)$ in the third simulation study. (a) The absorption image used as the reference in the error computations. (b) The reconstructed absorption image using the uniform mesh in figure 2(a) for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (c) The reconstructed absorption image using adaptive meshes based on theorem 2 for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (d) The reconstructed absorption image using adaptive meshes based on *a priori* error estimates (2.16) and (2.17) for the forward, and the uniform mesh in figure 2(b) for the inverse problem. (e) The reconstructed absorption image using adaptive meshes based on theorem 2 for the forward, and the adaptive mesh based on theorem 1 for the inverse problem. (f) The reconstructed absorption image using adaptive meshes based on *a priori* error estimates (2.16) and (2.17) for the forward, and the interpolation error estimate (2.14) for the inverse problem.

5. Conclusion

In this work, based on the error analysis presented in part I [9], we developed two new adaptive mesh generation algorithms, one for the forward and one for the inverse problem, which take into account the interdependence between the solutions of the two problems. We have also presented the computational complexity of the presented adaptive mesh generation algorithms. Our numerical experiments provided a verification of theorems 1 and 2 and showed that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems. Conventional error estimates do not include domain-specific factors. As a result, the adaptive mesh generation algorithms based on conventional error estimates (2.16)–(2.17) and (2.14) may lead to high errors in reconstructed optical images as demonstrated in our numerical experiments.

We finally note that the adaptive mesh generation algorithms introduced in this paper can be adapted for the forward and inverse problems of similar inverse parameter estimation problems, such as electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, microwave imaging etc.

Acknowledgments

This material is based upon work supported by NSF-BES-0353160, ONR-N00014-04-1-0694 and US Army Medical Research-W81XWH-04-1-0559. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Various portions of this research were supported by the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (Award Number EEC-9986821) and Rensselaer Polytechnic Institute.

Appendix. Solution of the model problem

In order to initialize the adaptive mesh for the solution of the forward problem (provided $D(\mathbf{x}) = \overline{D}$ and $\mu_a(\mathbf{x}) = \overline{\mu_a}$ are spatially constant), we use an analytical solution to compute the estimates of g_j and g_i^*. Below, we give the solution in 2D for (2.1). Under the same conditions, an analytical solution for the adjoint problem (2.3) can be obtained in a similar way.

First, we use the polar coordinates (ρ, θ) to rewrite (2.1):

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial g_j}{\partial\rho}\right) + \frac{1}{\rho}\frac{\partial}{\partial\theta}\left(\rho\frac{\partial g_j}{\partial\theta}\right) + K_\Omega^2 g_j = -\frac{4\pi}{\rho}\frac{\delta\left(\rho-\rho_s^j\right)\delta\left(\theta-\theta_s^j\right)}{\overline{D}},$$

where we consider an unbounded domain, model the point source located at $\left(\rho_s^j, \theta_s^j\right)$ by the Dirac-delta function $4\pi\delta\left(\rho-\rho_s^j\right)\delta\left(\theta-\theta_s^j\right)/\rho$ and $K_\Omega^2 = -(\overline{\mu_a}c+\mathrm{i}\omega)/c\overline{D}$. Then, the solution g_j at (ρ, θ) due to the point source located at $\left(\rho_s^j, \theta_s^j\right)$ is given by [11]

$$g_j\left(\rho, \rho_s^j; \theta, \theta_s^j\right) = \frac{4}{\overline{D}\pi}\left\{\frac{1}{2}I_0(k_\Omega\rho_<)K_0(k_\Omega\rho_>) + \sum_{m=1}^{\infty}\cos[m(\theta-\theta_s)]I_m(k_\Omega\rho_<)K_m(k_\Omega\rho_>)\right\},$$

where $\rho_<$ means the smaller of ρ and ρ_s^j, $\rho_>$ means the greater of ρ and ρ_s^j, I_m and K_m are the modified Bessel functions of the first and second kind, respectively [1] and $k_\Omega = \sqrt{-K_\Omega^2}$.

The solution of the problem in 3D can be derived in a similar manner [11, 17].

References

[1] Abramowitz M and Stegun I A 1968 *Handbook of Mathematical Functions* (New York: Dover)

[2] Arridge S R 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41–93

[3] Arridge S R, Kaipio J P, Kolehmainen V, Schweiger M, Somersalo E, Tarvainen T and Vauhkonen M 2006 Approximation errors and model reduction with an application in optical diffusion tomography *Inverse Problems* **22** 175–95

[4] Brankov J G, Yang Y and Wernick M N 2004 Tomographic image reconstruction based on a content-adaptive mesh model *IEEE Trans. Med. Imaging* **23** 202–12

[5] Brenner S C and Scott L R 2002 *The Mathematical Theory of Finite Element Methods* (Berlin: Springer)

[6] Eppstein M J, Dougherty D E, Hawrysz D J and Sevick-Muraca E M 2001 Three-dimensional Bayesian optical image reconstruction with domain decomposition *IEEE Trans. Med. Imaging* **20** 147–63

[7] Gu X, Xu Y and Jiang H 2003 Mesh-based enhancement schemes in diffuse optical tomography *Med. Phys.* **30** 861–9

[8] Guven M, Yazici B, Intes X and Chance B 2003 An adaptive multigrid algorithm for region of interest diffuse optical tomography *Int. Conf. in Image Processing Proc. of IEEE* **2** 823–6

[9] Guven M, Yazici B, Kwon K, Giladi E and Intes X 2007 Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: I *Inverse Problems* **23** 1115–33

[10] Huang M and Zhu Q 2004 Dual-mesh optical tomography reconstruction method with a depth correction that uses *a priori* ultrasound information *Appl. Opt.* **43** 1654–62

[11] Jackson J D 1962 *Classical Electrodynamics* (New York: Wiley)

[12] Joshi A, Bangerth W and Sevick-Muraca E M 2004 Adaptive finite element based tomography for fluorescence optical imaging in tissue *Opt. Exp.* **12** 5402–17

[13] Kress R 1999 Linear integral equations *Applied Mathematical Sciences* vol 82 2nd edn (Berlin: Springer)

[14] Molinari M, Cox S J, Blott B H and Daniell G J 2001 Adaptive mesh refinement techniques for electrical impedance tomography *Physiol. Meas.* **22** 91–6

[15] Rivara M C 1984 Mesh refinement processes based on the generalized bisection of simplices *SIAM J. Numer. Anal.* **21** 604–13

[16] Torregrossa M, Zint C V and Poulet P 2002 Image reconstruction in optical tomography: mesh influence *4th Int. Workshop: Computational Problems of Electrical Engineering* pp 183–6

[17] Walker S, Boas D A and Gratton E 1998 Photon density waves scattered from cylindrical inhomogeneities: theory and experiments *Appl. Opt.* **37** 1935–44

Adaptive mesh generation for diffuse optical tomography

(Invited Paper)

Murat Guven, Birsen Yazici
Electrical, Computer, and Systems
Engineering Department
Rensselaer Polytechnic Institute
Troy, New York 12180
Email: yazici@ecse.rpi.edu

Eldar Giladi
Helicos BioSciences Corporation,
Cambridge, Massachusetts 02139

Xavier Intes
Department of Biomedical Engineering
Rensselaer Polytechnic Institute
Troy, New York 12180

Abstract— **In Diffuse Optical Tomography (DOT), the discretization error in the numerical solutions of the forward and inverse problems results in error in the reconstructed optical images. In this work, based on the analysis presented in [5], we present two theorems that constitute the basis for adaptive mesh generation for the forward and inverse DOT problems. The proposed discretization schemes lead to adaptively refined composite meshes that yield the desired level of imaging accuracy while reducing the size of the discretized forward and inverse problems. Our numerical experiments validate the error estimates developed in [5] and show that the new adaptive mesh generation algorithms improve the accuracy of the reconstructed optical images.**

I. INTRODUCTION

Imaging in Diffuse Optical Tomography (DOT) is composed of two interdependent stages which seek solutions to the forward and inverse problems. The forward problem deals with describing the Near Infrared (NIR) light propagation and the inverse problem is concerned with the estimation of the unknown optical parameters from boundary measurements [1].

Numerical approaches in solving the forward and inverse problems in Diffuse Optical Tomography (DOT) poses a tradeoff between computational efficiency and imaging accuracy. This tradeoff is a direct consequence of the discretization of the forward and inverse problems [5], [1] and the size of the resulting discrete forward and inverse problems. Attempting to minimize the discretization error in the solutions of both problems separately results in increased size of the discrete forward and inverse problems. Thus, it is important to understand the relationship between the discretization error and the resulting error in the solution of the inverse problem.

In [5], we presented an error analysis that shows the effect of discretization of the forward and inverse problems on the accuracy of DOT imaging. In this work, based on the two error bounds provided by the error analysis in [5], we present an adaptive discretization scheme for the forward and inverse problems, respectively. We remark that the mesh refinement criterion for each problem comprises the discretization error in the corresponding problem solution, scaled spatially by the solutions of both problems. Thus, the proposed adaptive mesh generation algorithms address the interdependence between the solutions of the forward and inverse problems and take into account the orientation of the source-detectors and the absorptive perturbations. This makes the adaptive discretization algorithms presented in this paper different from the previous approaches [4], [9], [3], [7], [8] (see [5], [6] for an extensive literature survey). The simulation experiments validate the implications of our error analysis and show that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems.

The outline of this paper is as follows: In Section 2, we give a brief overview of the forward and inverse DOT problems and recall the two theorems presented in [5] which summarize the impact of discretization on the accuracy of the reconstructed optical images. Finally we discuss the adaptive mesh generation algorithms followed by our experimental results and the conclusion section.

TABLE I

INDEX OF NOTATION.

Notation	Explanation
$\|\|f(\mathbf{x})\|\|_0$	The $L^2(\Omega)$ (or $H^0(\Omega)$) norm of $f(\mathbf{x})$
$\|\|f(\mathbf{x})\|\|_p$	The $H^p(\Omega)$ norm of $f(\mathbf{x})$
$\|\|f(\mathbf{x})\|\|_{L^p(\Omega)}$	The $L^p(\Omega)$ norm of $f(\mathbf{x})$
$\|\|f(\mathbf{x})\|\|_{0,m}$	The L^2 (or H^0) norm of $f(\mathbf{x})$ over the mth finite element Ω_m
$\|\|f(\mathbf{x})\|\|_{p,m}$	The H^p norm of $f(\mathbf{x})$ on the mth finite element Ω_m
$\overline{f(\mathbf{x})}$	The complex conjugate of $f(\mathbf{x})$

II. OVERVIEW

In this section, we first briefly define the forward and inverse problems in DOT. Next, we state Theorems 1 and 2 presented in [5] to recall the effect of the discretization of the forward and inverse problems on the accuracy of optical absorption image reconstruction. We refer to Table I for the explanation of the notation associated with functions and their norms.

A. Forward and inverse problems in DOT

We consider the following boundary value problem to model the near infrared light propagation in a bounded domain $\Omega \subset$

R^3, with Lipschitz boundary [2]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_j(\mathbf{x}) + (\mu_a(\mathbf{x}) + \frac{i\omega}{c})g_j(\mathbf{x}) = Q_j(\mathbf{x}) \;\; \mathbf{x} \in \Omega, \quad (1)$$

$$g_j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_j}{\partial n}(\mathbf{x}) = 0 \;\; \mathbf{x} \in \partial\Omega, \quad (2)$$

where $g_j(\mathbf{x})$ is the photon density, $Q_j(\mathbf{x})$ is the point source located at the source position $\mathbf{x}_s^j$, $D(\mathbf{x})$ is the diffusion coefficient, $\mu_a(\mathbf{x})$ is the absorption coefficient, ω is the modulation frequency of the source, c is the speed of the light, a is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial \cdot /\partial n$ denotes the directional derivative along the unit normal vector $\mathbf{n}$ on the boundary. The boundary value problem (1)-(2) constitutes the forward problem in DOT together with the associated adjoint problem [1], [5]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^*(\mathbf{x}) + (\mu_a(\mathbf{x}) - \frac{i\omega}{c})g_i^*(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega, \quad (3)$$

$$g_i^*(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_i^*}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega, \quad (4)$$

where $Q_i^*(\mathbf{x})$ is the adjoint source located at the detector position $\mathbf{x}_d^i$. Note that we approximate the point source Q_j in (1) and the adjoint source Q_i^* in (4) by Gaussian functions with sufficiently low variance, whose centers are located at $\mathbf{x}_s^j$ and $\mathbf{x}_d^i$, respectively [5].

In this work, we focus on the estimation of the absorption coefficient and consider an iterative algorithm based on repetitive linearization of the inverse problem using first order Born approximation. Using a zeroth order Tikhonov regularization to address the illposedness, the inverse problem at each iteration reads:

$$\begin{aligned}\gamma(\mathbf{x}) = A_a^*\Gamma_{i,j} &= (A_a^*A_a + \lambda I)\alpha^\lambda = K\alpha^\lambda, & (5)\\ &= \int_\Omega \kappa(\mathbf{x},\acute{\mathbf{x}})\alpha^\lambda(\acute{\mathbf{x}})d\acute{\mathbf{x}} + \lambda\alpha^\lambda(\mathbf{x}) & (6)\end{aligned}$$

where $\Gamma_{i,j}$ is the differential measurement at the ith detector due to the jth source as a result of the small perturbation $\alpha(\mathbf{x})$ on the background absorption coefficient $\mu_a(\mathbf{x})$. In (6), $\kappa(\mathbf{x},\acute{\mathbf{x}})$ is the kernel of the integral equation, given by

$$\kappa(\mathbf{x},\acute{\mathbf{x}}) = \sum_{i,j}^{N_d,N_s} H^*(\mathbf{x};\mathbf{x}_s^j,\mathbf{x}_d^i)H(\mathbf{x}_s^j,\mathbf{x}_d^i;\acute{\mathbf{x}}), \quad (7)$$

where $H(\mathbf{x}_s^j,\mathbf{x}_d^i;\acute{\mathbf{x}}) = -\overline{g^*(\acute{\mathbf{x}},\mathbf{x}_d^i)}g(\acute{\mathbf{x}},\mathbf{x}_s^j)$ is the kernel of the integral operator $A_a : L^\infty(\Omega) \to C^{N_d\times N_s}$ and $H^*(\mathbf{x};\mathbf{x}_s^j,\mathbf{x}_d^i) = -\overline{g^*(\mathbf{x},\mathbf{x}_d^i)}g(\mathbf{x},\mathbf{x}_s^j)$ is the kernel of the adjoint operator $A_a^* : C^{N_d\times N_s} \to L^1(\Omega)$ [6]. For the rest of the paper, we will denote $L^\infty(\Omega)$ and $L^1(\Omega)$ by X and Y, respectively.

Below we summarize the two theorems of [5] and provide the error estimates which will be used in the design of adaptive meshes for the discretization of the forward and inverse DOT problems.

B. Effect of inverse problem discretization

Consider the discretization of the inverse problem (5) by projecting it onto a finite dimensional subspace Y_n, using the collocation method [5]:

$$P_nK\alpha_n^\lambda = P_n\gamma, \quad (8)$$

where $P_n : Y \to Y_n$ is the projection operator associated with the collocation method with piecewise linear Lagrange basis functions $L_p(\mathbf{x})$ [5] such that $\alpha_n^\lambda(\mathbf{x})$, $\mathbf{x} \in \Omega$, is approximated on a set $\{\Omega_m\}$ of finite elements for $m = 1, \cdots, N_\Delta$, $\bigcup_m^{N_\Delta} \Omega_m = \Omega$ as follows:

$$\alpha_n^\lambda(\mathbf{x}) = \sum_{p=1}^{n} a_pL_p(\mathbf{x}). \quad (9)$$

Let ψ be the interpolant of α^λ [2] and assume that $\alpha^\lambda \in H^1(\Omega)$. Then, the interpolation error $e_\alpha = \alpha^\lambda - \psi$ is bounded by

$$\|e_\alpha\|_{0,m} \le C\|\alpha^\lambda\|_{1,m}h_m, \quad (10)$$

where C is a positive constant and h_m is the diameter of the smallest ball that contains the mth element. Then,

Theorem 1:

The error between the solution α^λ of (5) and the solution α_n^λ of (8) is bounded by $\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)}$:

$$\begin{aligned} &\le \; C\sqrt{V_\Omega}\|I - T_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\|\alpha^\lambda\|_{1,m}h_m \\ &+ \; \frac{C}{\lambda}\|T_n\|_{Y\to X_n}\max_{i,j}\|g_i^*g_j\|_{L^1(\Omega)} \\ &\times \; \sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s}\|g_i^*g_j\|_{0,m}\|\alpha^\lambda\|_{1,m}h_m, \end{aligned}$$

where C is a positive constant, V_Ω is the volume of Ω, $T_n : Y \to X_n$ is a uniformly bounded operator given by $T_n = [I + \frac{1}{\lambda}P_nA_a^*A_a]^{-1}P_n$, and g_j, g_i^* are the solutions of the variational formulations of the boundary value problems (1)-(2) and (3)-(4), respectively [5].

Proof: See [5].

C. Effect of forward problem discretization

Assume that $D(\mathbf{x}), \mu_a(\mathbf{x}) \in C^2(\Omega)$. Noting that $Q_j, Q_i^* \in H^1(\Omega)$, the solutions g_j, g_i^* of the variational formulations of the boundary value problems (1)-(2) and (3)-(4) satisfy [5]

$$g_j(\mathbf{x}), g_i^*(\mathbf{x}) \in C^1(\Omega). \quad (11)$$

Let $G_j(\mathbf{x})$ and $G_i^*(\mathbf{x})$ be the finite element approximations to g_j and g_i^*, respectively, and let h_m^j and h_m^i be the diameter of the smallest ball that contains the mth element in the finite element solutions G_j and G_i^*, respectively. Then, a bound for the discretization error in the solutions G_j and G_i^* can be given by

$$\|g_j - G_j\|_{0,m} \le C\|g_j\|_{1,m}h_m^j, \quad (12)$$

$$\|g_i^* - G_i^*\|_{0,m} \le C\|g_i^*\|_{1,m}h_m^i, \quad (13)$$

where C is a positive constant.

Consider the inverse problem

$$\tilde{K}\tilde{\alpha}^\lambda = \tilde{\gamma}, \quad (14)$$

where $\tilde{K}$ and $\tilde{\gamma}$ are the finite dimensional approximations to K and γ, obtained by substituting g_j and g_i^* in $H(\mathbf{x}_s^j, \mathbf{x}_d^i; \acute{\mathbf{x}})$ and $H^*(\mathbf{x}; \mathbf{x}_s^j, \mathbf{x}_d^i)$ by G_j and G_i^*, respectively. Then,

Theorem 2:

A bound for the error between the solution α^λ of (5) and the solution $\tilde{\alpha}^\lambda$ of (14) due to approximations $\tilde{K}$ and $\tilde{\gamma}$ is given by:

$$\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)} \leq \frac{C}{\lambda} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)}$$
$$\times \left(\sum_{m=1}^{N_\triangle} \sum_{i,j}^{N_d, N_s} \left(\ 2\|g_j \alpha^\lambda\|_{0,m} + \|\alpha\|_\infty \|g_j\|_{0,m} \ \right) \|g_i^*\|_{1,m} h_m^i \right.$$
$$\left. + \sum_{m=1}^{N_\triangle} \sum_{i,j}^{N_d, N_s} \left(\ 2\|g_i^* \alpha^\lambda\|_{0,m} + \|\alpha\|_\infty \|g_i^*\|_{0,m} \ \right) \|g_j\|_{1,m} h_m^j \right),$$

where C is a positive constant.

Proof: See [5].

We refer to [5] for further details of the discussion regarding the definition and discretization of the forward and inverse problems.

III. ADAPTIVE MESH GENERATION

The adaptive mesh generation is based on minimizing the error bounds in Theorems 1 and 2 and distributing the error bound evenly on each of the finite elements. For the details of the mesh generation, we refer to [6].

IV. NUMERICAL EXPERIMENTS

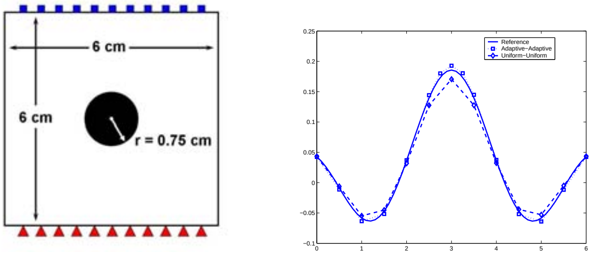
In the following, we present the results of our numerical experiments. Note that in all experiments, we use triangular finite elements with piecewise linear Lagrange basis functions. We apply Gaussian elimination method to solve the discrete forward problem resulting from the variational formulation [2] of the boundary value problems (1)-(2) and (3)-(4) [5]. For the inverse problem, we consider the discrete problem obtained by projecting (14) by collocation method:

$$P_n \tilde{K} \tilde{\alpha}_n^\lambda = P_n \tilde{\gamma}, \tag{15}$$

where the regularization parameter is set to $\lambda = 10^{-7}$, which is chosen based on experience.

A. Simulation Study

In this simulation study, we consider the geometry shown in Figure 1(a). We simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 61 nodes along x and y directions, where the refractive index mismatch parameter $a = 3$. 11 sources and 11 detectors are evenly spaced on the bottom and top edges of the square, respectively. The diffusion coefficient $D(\mathbf{x}) = 0.0410$ is assumed to be constant. The circular heterogeneity with absorption coefficient $\mu_a = 0.2$ cm^{-1} is embedded in an optically homogeneous background with $\mu_a = 0.04$ cm^{-1}.



(a) The optical domain and source-detector configuration for the simulation study.

(b) The solid, square, and diamond lines correspond to the cross-sectional cuts taken from the images shown in Figures 3(a), 3(d), and 3(b), respectively.

Fig. 1. **(a)** The setup used for the simulation study. The squares and triangles denote the detectors and sources, respectively. **(b)** The cross-sectional cuts taken from Figures 3(a), 3(b), and 3(d), along x direction at $y = 3$.

In order to obtain a series of absorption imaging problems using the same setup, we consider 5 values for the background absorption value. Then, for each imaging problem, we consider three mesh scenarios: Uniform mesh for both forward and inverse problems; adaptive mesh for the forward problem and uniform mesh for the inverse problem; and adaptive meshes for both forward and inverse problems. We refer to Table II for a brief outline of the first simulation study.

TABLE II

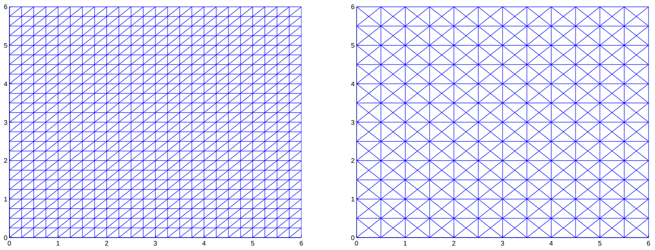
THE MESH SCENARIOS AND THE BACKGROUND μ_a VALUES IN THE SIMULATION STUDY.

Mesh (Forward)	**Mesh (Inverse)**	Background μ_a (cm^{-1})
Uniform	Uniform	0.032, 0.036, 0.040, 0.044, 0.050
Adaptive	Uniform	0.032, 0.036, 0.040, 0.044, 0.050
Adaptive	Adaptive	0.032, 0.036, 0.040, 0.044, 0.050

The uniform mesh used for the forward problem discretization has 625 nodes and is shown in Figure 2(a). The uniform mesh for the inverse problem has 313 nodes and is shown in Figure 2(b). We use the algorithms described in Section 3.1 and Remark 1(iii), and Section 3.3 Remark 2(iii) in [6] to generate the adaptive meshes for the forward and inverse problems, respectively. The number of nodes in each of the adaptive meshes used for the forward problem does not exceed 750. An example for the adaptive mesh generated for a source located at $(1.0, 0)$ is shown in Figure 2(c). The adaptive mesh for the inverse problem generated for the case where the background $\mu_a = 0.050$ cm^{-1} has 418 nodes and is shown in Figure 2(d).

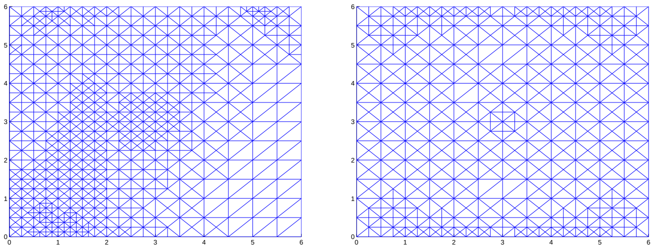
We consider the image reconstructed by using fine uniform meshes (61 × 61 nodes for the forward problem and 61 × 61 nodes for the inverse problem) as the reference image, which is assumed to possess no error due to discretization. We compute the error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each image reconstruction and tabulate the results in Table III. We see that the error in the images reconstructed by using uniform meshes for both forward and inverse problems is significantly reduced by the use of adaptively refined meshes. A similar behavior is observed for all choices of background absorption value.

We present image reconstructions in Figures 3 for the



(a) The uniform mesh with 625 nodes.

(b) The uniform mesh with 313 nodes.



(c) The adaptive mesh generated for the forward problem for the source located at (1.0,0): Background $\mu_a = 0.050$ cm^{-1}.

(d) The adaptive mesh generated for the inverse problem solution, with 418 nodes. Background $\mu_a = 0.050$ cm^{-1}.

Fig. 2. Examples of meshes used in the first simulation study.

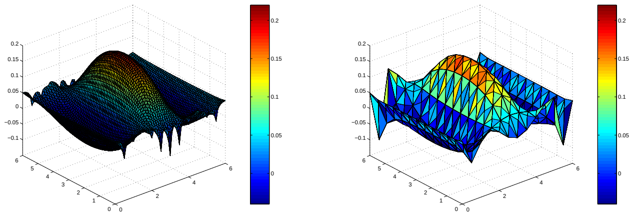
TABLE III

THE ERROR $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ FOR EACH EXPERIMENT DESCRIBED IN THE SIMULATION STUDY AND TABLE II. THE FIRST COLUMN SHOWS THE TYPE OF THE MESHES ("U" FOR UNIFORM, "A" FOR ADAPTIVE) USED IN THE FORWARD AND INVERSE PROBLEMS, RESPECTIVELY.

	Background μ_a:	0.032	0.036	0.040	0.044	0.050
U-U	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.233	0.256	0.277	0.293	0.301
A-U	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.124	0.114	0.117	0.121	0.128
A-A	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.104	0.099	0.099	0.100	0.101

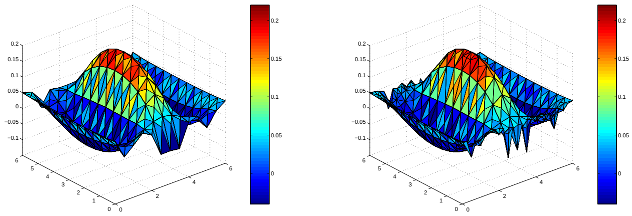
case, in which the background absorption value is equal to 0.050 cm^{-1}. Figure 3(a) displays the reference image used to compute the corresponding error values given in Table III. Figures 3(c) and 3(d) show that the optical heterogeneity is resolved better by using adaptive meshes as compared to the reconstructed image obtained by using uniform meshes, which is shown in Figure 3(b). These results are consistent with the error values given in Table III. Note that the number of nodes in the adaptive meshes is almost equal to the number of nodes that the uniform meshes have. In Figure 1(b), we show the cross-sectional views from the reconstructed images.

V. CONCLUSION



(a) The optical absorption image used as the reference for error computations.

(b) The reconstructed absorption image using the uniform mesh in Figure 2(a) for the forward, and the uniform mesh in Figure 2(b) for the inverse problem.



(c) The reconstructed absorption image using an adaptive mesh for the forward, and the uniform mesh in Figure 2(b) for the inverse problem.

(d) The reconstructed absorption image using an adaptive mesh for the forward, and the adaptive mesh in Figure 2(d) for the inverse problem.

Fig. 3. The results of the simulation study, with the background $\mu_a = 0.050$ cm^{-1}.

In this work, based on the error analysis presented in [5], we presented a verification of Theorems 1 and 2 and showed that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems. Conventional error estimates do not include domain specific factors. As a result, the adaptive mesh generation algorithms based on conventional error estimates (12)-(13) and (10) may lead to high errors in reconstructed optical images (see [6]).

We finally note that the adaptive mesh generation algorithms presented in this paper can be adapted for the forward and inverse problems of similar inverse parameter estimation problems, such as electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, microwave imaging etc.

REFERENCES

[1] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:R41–93, 1999.

[2] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer Verlag, 2002.

[3] X. Gu, Y. Xu, and H. Jiang. Mesh-based enhancement schemes in diffuse optical tomography. *Medical Physics*, 30(5):861–869, 2003.

[4] M. Guven, B. Yazici, X. Intes, and B. Chance. An adaptive multigrid algorithm for region of interest diffuse optical tomography. In *International Conference in Image Processing*, volume 2 of *Proc. of IEEE*, pages 823–826, 2003.

[5] M. Guven, B. Yazici, K. Kwon, E. Giladi, and X. Intes. Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: Part I. *Submitted to Inverse Problems.*

[6] M. Guven, B. Yazici, K. Kwon, E. Giladi, and X. Intes. Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: Part II. *Submitted to Inverse Problems.*

[7] M. Huang and Q. Zhu. Dual-mesh optical tomography reconstruction method with a depth correction that uses a priori ultrasound information. *Applied Optics*, 43(8):1654–1662, 2004.

[8] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca. Adaptive finite element based tomography for fluorescence optical imaging in tissue. *Optics Express*, 12(22):5402–5417, 2004.

[9] M. Torregrossa, C. V. Zint, and P. Poulet. Image reconstruction in optical tomography: mesh influence. In *IV-th International Workshop*, Computational Problems of Electrical Engineering, pages 183–186, 2002.